Area Exam: Theory and Practice of Reconfigurable Optical Networks

Matthew Nance Hall
Department of Computer and Information Science
University of Oregon
Fall 2020

CONTENTS

Abstract—**Reconfigurable optical networks have emerged as a promising technology to serve the fast-growing traffic produced by the digital society. In this area exam we consider reconfigurable optical networks and their interfaces to higher layers of the networking stack. To this end, we explore the challenges for implementing a vertically programmable network. First, we survey modeling work which is essential for efficiently utilizing reconfigurable optical networks given limited resources. Then, we discuss practical implementations for reconfigurable optical networks including hardware technologies and systems implementations. Finally, we explore exciting applications for future work in this field, including network simulation, measurement, traffic engineering, and cybersecurity.**

I. INTRODUCTION

We are amidst an explosive growth in communication traffic driven by data-centric applications related to business, science, social networking, and entertainment. Further, the rise of machine learning and artificial intelligence creates yet more demand for data-hungry applications. This trend affects both data center and wide-area networks. Researchers have responded by innovating at different layers of the network stack. For example, software-defined networking controllers that run traffic engineering applications are replacing ad-hoc routing [1], [2]. Software-based systems are also replacing hardware appliances such as load balancers [3], [4]. Programmable switches are replacing proprietary and vendor-specific switch APIs [5], [6]. Finally, even server NICs are being re-imagined as cloud network operators look to FPGA implementations as an avenue for scaling bandwidth [7]. Reconfigurable optical networks are the final frontier in this trend, where programmability is embedding itself deeper into the network stack.

Reconfigurable optical technologies are an incredibly exciting innovation. They enable adaptation of both the topology itself and of the network capacity on links. Such adaptations may be exploited by next-generation systems to improve performance and efficiency (*e.g.*, by making the network topology demand aware). For example, recent technologies based on free-space optics or optical circuit switches support very fast topology adaptations in data centers. Meanwhile, technologies based on reconfigurable optical add-drop multiplexers (ROADMs) can add or drop wavelengths carrying data channels from a transport fiber without the need to convert the signals to electronic signals and back. In both cases, operators need not carry out the entire bandwidth assignment and optical route planning during the initial deployment.

While research gains momentum in virtualizing the network stack, we are coming closer to a vertically programmable network. Ideally, in such an ecosystem, every aspect of connectivity in a network is programmable. A vertically programmable network yields unparalleled flexibility for routing, topology adaptation, and bandwidth assignment. However, the virtualization of the physical layer is uniquely challenging. With reconfigurable optical networks being a relatively new technology, the community is still discussing their conceptual fundamentals, benefits, and limitations. There is a massive disconnect, or *chasm*, between the optical communications layer and higher layers of the network stack. This chasm is at the core of challenges for vertically programmable networks.

This area exam is an up-to-date survey on emerging reconfigurable optical networks. While there are surveys with broad overviews of reconfigurable optical technologies in these settings (e.g., software-defined optical networks [8], routing and spectrum allocation [9], wavelength switching hardware architecture [10]), the functioning of such systems (e.g., reconfigurable metropolitan networks [11] and data center networks [12]) requires a full-stack perspective on optical networks. We address this requirement by presenting an *end-to-end perspective* on reconfigurable optical networks by (a) emphasizing the interdependence of optical technologies with algorithms and systems and (b) identifying the open challenges and future work at the intersection of optics, theory, algorithms, and systems communities.

This work is especially timely, as interest in dynamic optical layer networking technologies is gaining attention from the networked systems community. Upon reviewing the last five years of publications from five optical and systems network journals, we ran a clustering analysis to see how much overlap there has been between the two fields. Table I shows the journals, and their raw publication counts since 2015. Figure 1 shows the clustering analysis results for the ten largest clusters. Connections between papers are first-order citations. Based on the most massive cluster, 1, it appears there is a strong relationship between the Journal of Lightwave Technology (JLT) and the Journal of Optical Communications and Networking (JOCN) publications. Cluster 2 is mainly composed of IEEE Transactions on Communications (TCOM) papers, with a few IEEE/ACM Transactions on Networking (TON) papers. Cluster 3 shows a strong relationship between TON and IEEE Transactions on Network and Service Management (TNSM) papers. The three predominantly physical-layer journals (JLT, TCOM, and JOCN) appear together in cluster 4. Cluster 5 shows a relation between physical-layer and systems journals, TCOM, and TON. Clusters 6 and beyond are mostly singleton clusters, comprising predominantly one journal. Our analysis underscores a division between optical and higher layer publishing venues.

We address the apparent disconnect between networked systems and optical layer communities by surveying the most cutting-edge research in the last decade. Figure 2 shows the volume of work covered by this exam. The overwhelming majority of work cited is from the past five years. The papers covered here include works from the top networking conferences (*e.g.*, SIGCOMM, NSDI, CoNEXT) as well as optical networking and networked systems journals highlighted by our clustering analysis.

There already exist some excellent surveys on optical networks which at least partially cover reconfigurable aspects, both in the context of data centers [13]–[15] and (to a lesser extent) in the context of wide-area networks [16]. We extend these surveys while providing an up-to-date overview of the literature. Our paper aims to provide an understanding of the underlying fundamental issues and concepts in reconfigurable optical networks and identify commonalities and differences (spanning both data center and wide-area

Journal	Papers
IEEE-OSA Journal of Lightwave Technology	3,988
IEEE Transactions on Communications	2,703
IEEE-ACM Transactions on Networking	1,226
IEEE-OSA Journal of Optical Communications	817
IEEE Transactions on Network and Service Management	533

TABLE I: Papers published since 2015 in various networking journals.

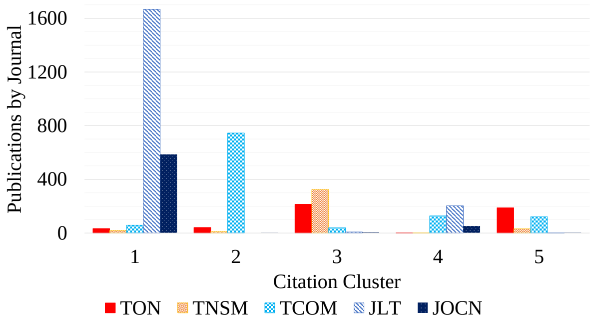

Fig. 1: Paper clusters among five networking and optical systems journals.

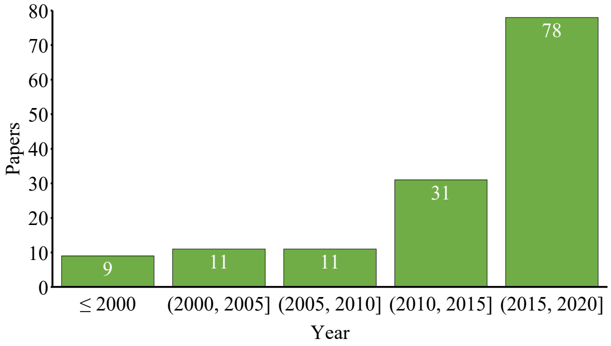

Fig. 2: Papers covered in this Area Exam by year.

networks). To this end, we proceed from theoretical models to practical technological constraints and implementations. We finish by exploring exciting applications for reconfigurable optical networks, including network simulation, measurement, traffic engineering, and cybersecurity.

The remainder of this paper is organized as follows. Section II defines the network architecture model for optical networks and its connection to the packet-switched network model. Section III illuminates modeling work in reconfigurable optical networks. Section IV is a broad overview of practical implementations of reconfigurable optical networks, first showcasing hardware technologies in Section IV-A. Then, Section IV-B showcases research on reconfigurable optical data center networks (DCNs) by highlighting DCN-specific hardware capabilities, algorithms, and systems implementations. Section IV-C looks directly at reconfigurable wide-area networks (WANs) with an emphasis on WAN-specific challenges in addition to algorithms and systems implementations. In section V we look at future work and applications for reconfigurable optical networks. Section VI concludes with the overarching open challenges for the field of reconfigurable optical networks spanning hardware, data centers, and wide-area networks.

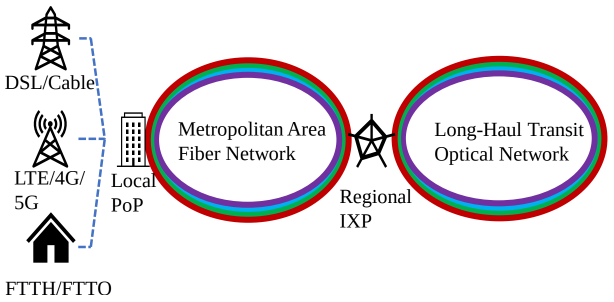

Fig. 3: Network architecture model, showing the connection between IP and optical layers.

II. NETWORK ARCHITECTURES

A. *IP-over-Optical Transport Network*

We discuss IP-over-Optical Transport Networks (IP-over-OTN) when referring to the network's IP and the Optical layers. In IP-over-OTN, hosts (*e.g.*, data centers, points-of-presence or PoPs, servers, *etc.*) connect to routers, and these routers are connected through the optical transport network (OTN) as shown in Figure 3. A node in the optical layer is an Optical Cross-Connect (OXC). An OXC transmits data on modulated light through the optical fiber. The modulated light is called a lambda, wavelength, or a circuit. The OXC can also act as a *relay* for other OXC nodes to transparently route wavelengths. When acting as a relay for remote hosts, an OXC provides *optical switching* capabilities, thus giving the network flexibility in choosing where to send transmitting lambdas over the OXC node.

IP-over-OTN networks are not new. However, they are built at a great cost. Historically network planners have engineered them to accommodate the worst-case expected demand by (1) over-provisioning of dense wavelength division multiplexing (DWDM) optical channels and (2) laying redundant fiber spans as a fail-safe for unexpected traffic surges. These surges could come from user behavior changes or failures elsewhere in the network that forces traffic onto a given path. Only recently have reconfigurable optical systems begun to gain attention in the data center and wide-area network settings. For more information about early IP-over-OTN, we defer to Bannister *et al.* [17] and references therein, where the authors present work on optimizing WDM networks for node placement, fiber placement, and wavelength allocation.

B. *Data Center Architecture*

Historically, data centers relied on packet-switched networks to connect their servers; however, as scale and demand increased, the cost to build and manage these packet-switched networks became too large. As a result of this change, new reconfigurable network topologies gained more attention from researches and large cloud providers. Many novel data center architectures with reconfigurable optical topologies have been proposed over the last decade. These architectures have in common that they reduce the static network provisioning requirements, thereby reducing the network's cost by presenting a means for bandwidth between hosts to change periodically.

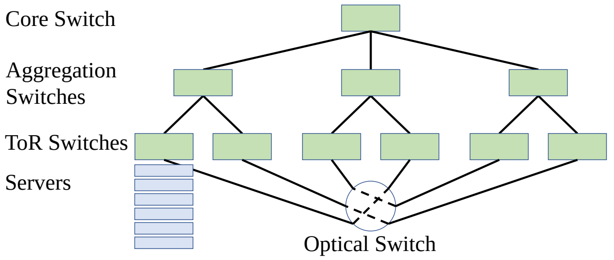

Fig. 4: Data center architecture proposed in C-Through [19]

Figure 4 shows one such example of a hybrid electrical-optical data center architecture. These architectures reduce cost and complexity via scheduling methods, which change bandwidth on optical paths in the data center. Various approaches have been demonstrated. Notable architectures employ fixed, and deterministic scheduling approaches [12], [18] or demand-aware changes that prioritize establishing optical paths between servers with mutual connectivity requests [19], [20]. Switching fabrics are also diverse for data center optical systems. These include fabrics based on nanosecond tunable lasers [21], digital micromirror devices (DMD) [22], and liquid crystal on silicon (LCOS) wavelength selective switches (WSS) [23].

III. THEORETICAL MODELS

A. Data centers

Momentum has been building for data centers to move to optically switched and electrical/optical hybrid networks. However, there is a general reluctance to walk away from the old paradigm of a packet-switched-only network due to the additional complexity of optical circuit switching (*e.g.*, the control plane management of optical circuits with shifting demand, and the variety of optical switching architectures available). Further, without a quantitative measure of *value-added* by optical switching over packet-switched-only, data center network operators are understandably reluctant to spend capital on an unvetted system.

To address the concerns surrounding complexity and value while raising awareness for the necessity of optically switched interconnects, researchers have constructed cost models to demonstrate the benefit of optical switching and hybrid architectures. Wang *et al.* [24] developed one such model. They conducted intra-DC traffic measurements, which consisted of mixed workloads (*e.g.*, Map-Reduce, MPI, and scientific applications). They then played the traces back in simulation, assuming that three optical circuits could be created and reconfigured between racks every 30 seconds. Their analysis in a data center with seven racks showed that rack-to-rack traffic over the packet-switched network could be reduced by 50% with circuit switching.

More theoretical modeling work is presented in Section IV-B, which considers practical implementations of reconfigurable data center networks. These papers also include evaluations of prototype systems, therefore we leave them out of this section on pure theoretical models.

B. Cost

Fiber infrastructure for wide-area networks is incredibly costly. Provisioning of fiber in the ground requires legal permitting processes through various governing bodies. As the length of the span grows beyond metropolitan areas, to connect cities or continents, the number of governing bodies with whom to acquire the legal rights to lay the fiber grows [25]. Then, keeping the fiber lit also incurs high cost; power requirements are a vital consideration for wide-area network providers [26]. Therefore, reliable cost models are necessary for deploying and managing wide-area networks. In this section, we look at cost modeling efforts particularly suited for reconfigurable optical networks.

An early study on the cost comparison of IP/WDM vs IP/OTN networks (in particular: European backbone networks) was conducted by Tsirilakis *et al.* in [27]. The IP/WDM network consists of core routers connected directly over point-to-point WDM links in their study. In contrast, the IP/OTN network connects the core routers through a reconfigurable optical backbone consisting of electro-optical cross-connects (OXCs) interconnected in a mesh WDM network.

Capacity planning is a core responsibility of a network operator in which they assess the needs of a backbone network based on the projected growth of network usage. Gerstel *et al.* [28] relates the capacity planning process in detail, which includes finding links that require more transponders and finding shared-risk-link-groups that need to be broken-up, among other things. They note that in this process, the IP and optical network topologies are historically optimized separately. They propose an improvement to the process via multi-layer optimization, considering the connection between IP and optical layers. They save 40 to 60% of the required transponders in the network with this multi-layer approach. The networks they looked at were Deutsche Telekom [29] and Telefonica Spain core networks. These authors' work provides a strong motivation for jointly optimizing IP and optical network layers and sharing of information between the two.

Papanikolaou *et. at.* [30] propose a cost model for joint multi-layer planning for optical networks. Their paper presents three network planning solutions; dual-plane network design, failure-driven network design, and integrated multi-layer survivable network design. They show that dual-plane and failure-driven designs over-provision the IP layer, leaving resources on the table that are only used if link failures occur. They show that integrated multi-layer survivable network design enables a significant reduction in CapEx and that the cost savings increases beyond dual plane and failure driven designs.

Cost models for evaluating colorless (C)-ROADM vs. colorless, directionless, contentionless (CDC)-ROADM network architectures are described by Kozdrowski *et al.* [31]. They show that for three regional optical networks (Germany, Poland, USA), CDC-ROADM based networks can offer 2 to 3× more aggregate capacity over C-ROADM based networks. They evaluate their model with uniform traffic matrices (TMs) and apply various scalar multipliers to the TM. Their model accounts for many optical hardware related constraints, including the number of available wavelengths and cost factors associated with manual-(re)configuration of C-ROADM elements. However, their model doesn't include an optical-reach constraint. They limit solver computation time to 20 hours and present the best feasible solution determined in that amount of time.

C. *Blocking Probability*

Blocking probability is a crucial metric for assessing the flexibility of an optical network. It is the probability that a request for an end-to-end lightpath in the network cannot be provisioned. Turkcu *et al.* [32] provides analytical probability models to predict the blocking probability in ROADM based networks with tunable transceivers and validate their models with simulation considering two types of ROADM architecture in their analysis, namely *share-per-node* and *share-per-link*. In share-per-link, each end of a link has a fixed number of transponders that can use it. In share-per-node, a node has a fixed set of transponders that may use any incident links. The authors show that low tunable range (4 to 8 channels, out of 32 possible) is sufficient for reducing blocking probability in two topologies, NSF Net (14 Nodes), and a ring topology with 14 nodes. As the tunable range moves beyond 8 and up to 32, there is little to no benefit for split-per-node and share-per-link architectures. As the load on the network increases, blocking probability increases, as well as the gap between blocking probability of split-per-node and split-per-link decreases.

D. *Service Velocity*

Service velocity refers to the speed with which operators may grow their network as demand for capacity grows. Woodward *et al.* [33] tackles the problem of increasing service velocity for WANs. In this context, they assume a network of colorless non-directional ROADMS (CN-ROADMs)[1], in which any incoming wavelength can be routed on any outgoing fiber. They claim that one of the largest impedances for network growth in these networks is the availability of *regenerators*. To solve this problem, they present three algorithms for determining regenerators' placement in a network as service demand grows. The algorithms are: locally aware, neighbor aware, and globally aware. Each algorithm essentially considers a broader scope of the network, which a node uses to determine if an additional regenerator is needed at the site at a particular time. They show, via Monte Carlo simulations, varying optical reach and traffic matrices. The broadest scope algorithm performs the best and allocates enough regenerators at the relevant sites without over-provisioning. This work shows that service velocity is improved with demand forecasting, enabling infrastructure to be placed to meet those projected demands.

E. *Competition*

Programmable and elastic optical networks can also work together with Network Function Virtualization (NFV) to offer lower-cost service-chaining to users. Optimal strategies have been demonstrated, with heuristic algorithms, to quickly find near-optimal solutions for users and service brokers by Chen *et al.* [34]. In their work, they take a game-theoretic approach to modeling the competition among service brokers—who complete offering the lowest cost optical routs and service chains, and between users—who compete to find the lowest cost and highest utility service chains among the brokers. They demonstrate both parties' strategies, which converge on low-latency service chain solutions with low blocking probability for optical paths.

Modeling *opportunity cost* of optically switched paths is explored by Zhang *et al.* [35]. In their work, they present an algorithm for quickly evaluating the opportunity cost of a wavelength-switched path. Given a request and a set of future requests, the opportunity cost for accommodating the initial request is the number of future requests blocked as a result of the accommodation. Thus, the network operator's goal is to minimize opportunity cost by permitting connections that interfere with the fewest future requests.

F. *Open Challenges*

The research literature is clear, that reconfigurable optical networks can save costs over static topologies. Further, with the right hardware in the network, we can construct networks with low blocking probability, i.e., networks that can use the full capability of wavelength switching to improve throughput. Also , by constructing networks with reconfigurability in mind, we can make sure that redundancies are in place that allow light paths to switch fiber paths more quickly, thereby giving even greater performance, especially in troubling scenarios such as fiber cuts.

Given these notions, we still currently lack a holistic modeling and simulation framework for prototyping and managing reconfigurable optical networks. Network simulators such as Mininet [36] exist for prototyping packet switched networks. Similarly, YATES [37] is a useful simulator for comparing traffic engineering applications. However, there is yet no simulation framework that integrates optical layer switching capability with the higher layers of the network stack. Such a framework could serve to be invaluable in designing next generation networks, whose underlying core topology is flexible and adaptable to change. In our ongoing work we seek to build such a system.

IV. PRACTICAL IMPLEMENTATIONS

A. *Enabling Hardware Technologies*

1) Wavelength Selective Switching (WSS): In contrast to packet-switched networks, optically circuit-switched systems operate at a more coarse granularity. The transmission of information over a circuit requires an end-to-end path for the communicating parties. Although packet switching has generally prevailed in today's Internet, recent research has revitalized the prospect of circuit switching for data centers and wide-area networks by illuminating areas in which flexible bandwidth benefits outweigh the start-up cost of circuit building.

---

[1] CN-ROADMs are also called CD-ROADMs in other papers. These both refer to the same ROADM architecture.

Technological advancements for optical hardware, primarily driven by physics and electrical engineering research, have been instrumental in making circuit-switched networks a viable model for data center networks. Among these technologies are low-cost/low-loss hardware architectures. Here we give a brief overview of technological advancements in this domain that have had the most significant impact on networked systems.

Kachris *et al.* [38] have an in-depth look at optical switching architectures in data centers from 2012. In their survey, they primarily look at competing *data center* architectures and switch models. In this section, we choose to focus instead on those architectures' physical manifestations (*i.e.*, the base components that make them up).

Polymer waveguides are a low-cost architecture for optical circuit switches. These have been fabricated and studied in depth over the last 20 years, including work by Taboada *et al.* [39] in 1999, Yeniay *et al.* [40] in 2004, and Felipe *et al.* [41] in 2018. Early implementations such as Taboada *et al.* [39] showed fabrication techniques for simple polymer waveguide taps. Multiple waveguide taps can be combined to form an Array Waveguide Grating (AWG), and the signals traversing the AWGs can then be blocked or unblocked to create an optical circuit switch. A major inhibitor of the polymer waveguide architecture was signal-loss, which was as high as 0.2 dB/cm until Yeniay *et al.* [40] discovered an improvement on the state-of-the-art with ultra low-loss waveguides in 2004. Their waveguides, made with fluorocarbons, have 4× less loss (0.05 dB/cm) than the next best waveguides at the time, made from hydrocarbons. Felipe *et al.* [41] demonstrate the effectiveness of a polymer waveguide-based switching architecture for reconfiguring groups of optical flows of up to 1 Tbps, proving that that AWG is a viable and competitive switching architecture for data centers. More recently, in 2020, AWGs were demonstrated to work in conjunction with nanosecond tunable transmitters to create flat topologies, significantly reducing power consumption for data center networks due to the passive—no power required—nature AWGs [42].

Microelectromechanical Systems (MEMs), introduced by Toshiyoshi *et al.* [43] in 1996, offered a lower-loss and more flexible alternative to polymer waveguide systems of the day. These MEMs devices are made up of small mirrors, which can be triggered between states (*i.e.*, *on* and *off*). Therefore, in a MEMs system light is *reflected* rather than *guided* (as in the polymer waveguide systems). This distinction between reflection and guiding implies generally slower switching speeds for MEMs based systems, as the mirror must be physically turned to steer light out of the desired switch-port. Despite this limitation, MEMs systems evolved to be competitive with polymer waveguides in modern systems. Data center solutions leveraging MEMs based switches include Helios [44].

Liquid Crystal on Silicon (LCOS) was demonstrated as another viable optical switching architecture by Baxter *et al.* [45] in 2006. In LCOS switches, multiple frequencies of light, which have been multiplexed together, are spatially separated and guided to a liquid crystal on silicon array. Each cell in the array corresponds to an input frequency, and the output port for that frequency is determined by applying a specific voltage to the silica in the cell. Switches based on the LCOS architecture are commercially available and are recognized as a key enabler for reconfigurable optical networks [23].

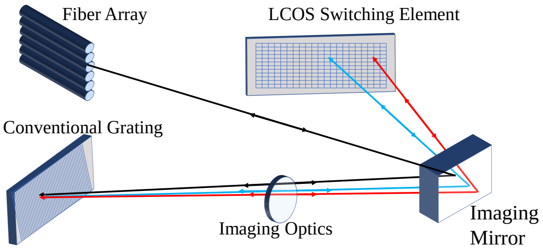

Fig. 5: Liquid crystal on silicon wavelength selective switch.

Figure 5 shows an abstracted LCOS WSS. Multiplexed optical signals from an array of fiber enter the system from a fiber array. These signals are directed to a conventional diffraction grating where the different colors or light are separated from each signal. These colors are then projected onto a unique position in the LCOS switching element. The voltage on the cell in the switching element determines which output fiber the channel will leave through. From there, the signal travels back through the system and into a different fiber in the array.

2) ROADMs: Reconfigurable add-drop multiplexers, or ROADMs, are an integral component of IP-over-OTN networks. These devices have evolved over the years to provide greater functionality and flexibility to optical transport network operators. We briefly describe the evolution of ROADM architectures. Figure 6 shows a broadcast and select ROADM architecture. Please refer to [10] for more information about ROADM architectures.

Colorless (C). Early ROADMs were effectively programmable wavelength *splitter-and-blockers*, or *broadcast-and-select* devices. A wavelength splitter-and-blocker can be placed before an IP-layer switch. If the switch is intended to *add/drop* a wavelength (*i.e.*, transceive data on it), then the blocker prohibits light on the upstream path and enables light on the path to the switch. These splitter-and-blocker systems are better known as Colorless, or C-ROADMs, as the *splitter-and-blocker* architecture is independent of any specific frequency of light. To receive the maximum benefit from C-ROADMs, operators should deploy their networks with tunable transceivers as they allow more flexibility for the end hosts when connecting to remote hosts.

Colorless, Directionless (CD). The CD-ROADMs extend the architecture of C-ROADMs by pairing multiple C-ROADMs together in the same unit to allow for a wave to travel in one of many directions. One shortfall of this architecture is that the drop ports from each direction are fixed, and therefore if all of the drop ports are used from one direction, the remaining points from other directions cannot be used. Due to the limitation of drop ports in different directions, the CD architecture is not *contentionless*.

Colorless, Directionless, Contentionless (CDC). The CDC-ROADM solves the contention problem by providing a shared add/drop port for each direction of the ROADM. This allows contentionless reconfiguration of the ROADM as any drop-

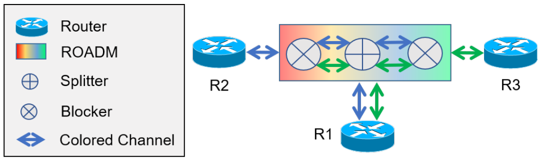

Fig. 6: Broadcast and Select colorful ROADM. The add/drop node, R1, has ports for two optical channels. These channels are directed at the ROADM. The ROADM uses a splitter to *broadcast* the channels onto two outbound ports, where a wavelength blocker *selects* the appropriate channel for the next router.

signal is routed to a common port regardless of the direction from which the wave begins/terminates.

Colorless, Directionless, Contentionless w. Flexible Grid (CDC-F). Flexi-grid, or elastic optical networks, are networks carrying optical channels with non-uniform grid alignment. This contrasts with a fixed-grid network, where different wavelengths are spaced with a fixed distance (*e.g.*, 50 GHz spacing). Wideband spacing allows signals to travel farther before becoming incoherent due to chromatic dispersion. Thus, CDC-Flex or CDC-F ROADMs enable the reconfiguration of wavelengths with heterogeneous grid alignments. These are most useful for wide area networks, with combinations of subsea and terrestrial circuits.

3) Bandwidth-variable Transponders: Before we discuss bandwidth-variable transponders, we must first take a moment to illuminate a common concept to all physical communications systems, not only optical fiber. This concept is modulation formats. Modulation formats determine the number of binary bits that a signal carries in one *symbol*. Two parties, a sender and receiver, agree on a symbol rate (baud), which determines a clock-speed to which the receiver is tuned when it interprets a symbol from the sender. The simplest modulation format is on-off keying (OOK), which transmits one-bit-per-symbol. In OOK, the symbol is sent via a high or low power level, as shown in figure 7A. A higher-order modulation technique is Quadrature Phase Shift Keying (QPSK), in which the symbol is a sinusoidal wave whose phase-offset relates the symbol. In this QPSK, there are four phase shifts agreed upon by the communicating parties, and therefore the system achieves two bits per symbol, or two baud, seen in figure 7B. A constellation diagram for QPSK is shown in figure 7C. As modulations become more complex, it is more useful to visualize them in the phase plane shown by their constellation diagram. Higher-order modulation formats are of the type, N-Quadrature Amplitude Modulation (QAM) techniques (Figure 7D), and these permit $log_2(N)$ bits per symbol[2]. In QAM, the symbol is denoted by phase and amplitude changes. Figure 7D shows an example of a constellation diagram for 16-QAM modulation, which offers 4 bits per symbol, or twice the baud of QPSK.

Fiber optic communications are subject to noise. The noise level is Signal to Noise Ratio (SNR), and this metric determines the highest possible modulation format. In turn, the modulation format yields a potential capacity (Gbps) for an optical channel. For example, in [46], the authors claim that SNR of just 6 dB is sufficient to carry a 100 Gbps signal, while a circuit with an SNR of 13 dB can transmit 200 Gbps.

Bandwidth Variable Transponders (BVTs) [47] have recently proven to have significant applications for wide-area networks. These devices are programmable, allowing for the operator to choose from two or more different modulation formats, baud rates, and the number of subcarriers when operating an optical circuit. For example, the same transponder may be used for high-capacity/short-reach transmission (16-QAM or greater) or lower-capacity/longer-reach transmission (*e.g.*, QPSK). Higher modulation formats offer higher data rates. They are also more sensitive to the optical SNR, which decreases in a step-wise manner with distance, as illustrated in Figure 8. BVTs enable network operators to meet the ever-growing demand in backbone traffic by increasing optical circuits' spectral efficiency.

Low spectrum utilization, or waste, can be an issue for BVT circuits. For example, a BVT configured for a low-modulation circuit such as QPSK instead of 16-QAM has a potential for untapped bandwidth. Sambo *et al.* [48] introduced an improvement to the BVT architecture, known as Sliceable-BVT (S-BVT), which addresses this issue. They describe an architecture that allows a transponder to propagate numerous BVT channels simultaneously. Channels in the S-BVT architecture are sliceable in that they can adapt to offer higher or lower modulation in any number of the given subchannels.

4) Silicon Photonics: Various materials (*e.g.*, GaAs, Si, SiGe) can be used to make photonics hardware required for data transmission. These devices include photodetectors, modulators, amplifiers, waveguides, and others. Silicon is the preferred material for these devices due to its low cost. However, there are challenges to manufacturing these silicon devices, such as optical power loss and free carrier absorption. Other materials, notably GaAs, have better properties for propagating light; however, GaAs is more costly to manufacture. Despite these challenges, research into efficient and quality transmission using silicon-based photonic devices has boomed in the last decade. Early advances were made towards silicon photonics in the 80s, particularly for waveguides, which are the basis for circuit switches and multiplexers. Today, silicon photonics is an integral part of almost all optical hardware, including lasers, photodetectors, modulators, and amplifiers. Costs are falling for optical hardware, thus enabling network operators to deploy newer technology into their systems at a more advanced pace as the devices' quality and guarantees have continued to improve. For more information on silicon photonics, see the survey by Thomson *et al.* [49].

5) Open Challenges: The development of hardware for reconfigurable optical networking is a burgeoning field in engineering and research. While CDC-F ROADMs exist today, they are costly to produce, and their capabilities are found lacking. In particular, the benefit of integrating CDC-F ROADMs with optical transport networks is limited by cascading fiber impairments, signal loss at WSS modules, and wavelength and fiber collision [50]. We expect silicon photonics to bring down the cost of transport hardware,

---

[2] where N is generally a power of 2

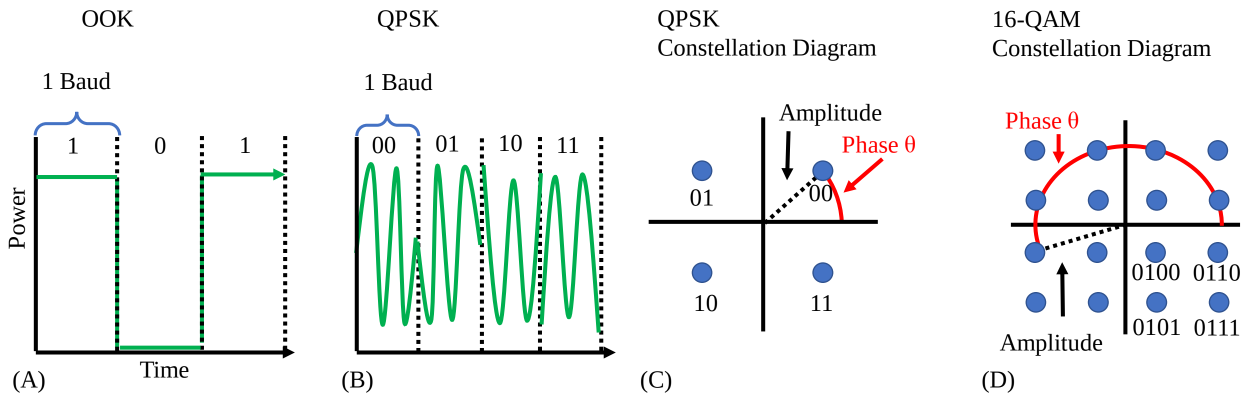

Fig. 7: Modulation examples of on-off keying, quadrature phase shift keying (QPSK), quadrature amplitude modulation (QAM), and constellation diagrams for QPSK and 16-QAM.

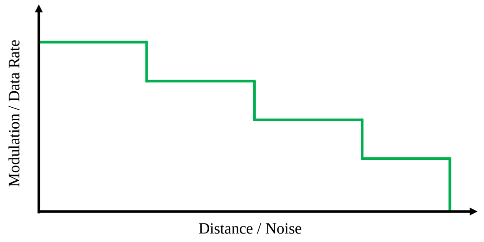

Fig. 8: Trade off between modulation/data rate and distance/noise with BVT.

thereby increasing access to such devices and lowering entry barriers for research and development.

Another great challenge here is that most optical networking components are invisible to higher layers of the networking stack. This consequence is a result of the passive nature of the devices. Unfortunately, it implies difficulty for efficiently managing and programmatically reconfiguring such devices. More critically, it creates a hard disconnect between the logical network topology and the physical, optical-wavelength topology. This means that it is nearly impossible to map optical networks with active measurement strategies.

B. Optically Reconfigurable Data Centers

A key challenge for data centers is to optimize the utilization of the data center network (DCN). In a DCN, many different services are running and competing for shared bandwidth. Communication patterns between top-of-rack (ToR) switches vary with the underlying applications that are running (*e.g.*, map-reduce, video stream processing, physics simulations, *etc.*). Thus, as future applications and user's needs change, it is challenging to predict where bandwidth will be needed.

Static and *reconfigurable* network solutions have been posed by research and industry to address this challenge. There is an assumption that the connectivity graph of the network cannot change in static network solutions. These solutions also assume fixed capacity (or bandwidth) on links. In reconfigurable network solutions, by contrast, these assumptions regarding connectivity and bandwidth are relaxed. Servers and switches (collectively referred to as nodes) may connect some subset of the other nodes in the network, and the nodes to which they are adjacent may change over time. Further, the bandwidth of a connection may also change over time.

Under the assumption of a static physical topology, different network architectures and best practices have been established. Some of these architectures include Clos, fat-tree, and torus topologies. Best practices include (over-)provisioning all links such that the expected utilization is a small fraction of the total bandwidth for all connections. These solutions can incur high cabling costs and are inefficient.

Reconfigurable network solutions circumvent the limitations of the static network solutions by reducing cabling costs or reducing the need to over-provision links. The flexibility of light primarily empowers these reconfigurable solutions. Some of these flexibilities include the steering of light (*e.g.*, with MEMs or polymer waveguides) and the high capacity of fiber-optics as a medium (*e.g.*, dense wavelength division multiplexing, or DWDM, enables transmitting $O(Tb/s)$ on a single fiber).

In this section, we illuminate efforts to improve DCNs with reconfigurable optics. Related surveys on this subject include Foerster *et al.* [13] and Lu *et al.* [15]. We divide the state of reconfigurable optical DCNs into *technology*, *algorithms*, and *systems*. In technology, we supplement the discussion from section IV-A with hardware capabilities that currently exist only for DCNs. Such features include free-space optics and sub-second switching. Next, we highlight cost modeling research, whose goal is to derive formal estimates or guarantees on the benefit of reconfigurable optical networks over static topologies for DCNs. Then, we survey the relevant algorithms for managing and optimizing reconfigurable optical networks in the data center and some systems implementations that leverage those algorithms.

1) DCN-specific Technologies: Innovations in reconfigurable optical networks are enabled by hardware's evolution, as discussed in section II. There is a subset of innovations that are well-suited for data centers only. These are *free-space optics* and *sub-second switching*. Although we have separated these below, there may be overlaps between free-space optics and sub-second switching systems as well.

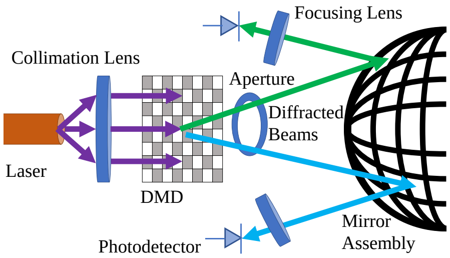

Fig. 9: Free-space optics switching architecture for data centers [52]

Free-space Optics. In free-space optics systems, light propagates through the air from one transceiver to another. Free-space optics enables operators to reduce their network's complexity (a function of cabling cost). These closed environments and their highly variable nature of intra-data center traffic make such solutions appealing. Recent works such as Firefly [51] have demonstrated that free space optics are capable of reducing latency for time-sensitive applications by routing high-volume/low-priority traffic over the wireless optical network while persistently serving low-volume/high-priority traffic on a packet-switched network. High fan-out (1-to-thousands) for free-space optics is enabled with DMDs, or Dense Micro-mirror Devices, as shown by ProjecToR [52]. The DMDs are placed near Top-of-Rack (ToR) switches and pair with disco-balls, fixed to the ceiling above the racks. The DMD is programmed to target a specific mirror on the disco-ball, guiding the light to another ToR in the data center. Figure 9 illustrates the main properties of the free space optics deployment proposed in [52]. The deployment and operation of a free-space optics data center are fraught with unique challenges, particularly for keeping the air clear between transceivers and DMDs. Any particulate matter that the light comes into contact with can severely degrade performance and cause link failures should they persist. This phenomenon is known as atmospheric attenuation [53].

Sub-second Switching. In data centers, distances are short between hosts, and therefore optical signals do not lose their strength to such a degree that mid-line devices such as amplifiers are necessary. Therefore, applications can benefit from all of the agility of optical layer devices without accounting for physical-layer impairments, which can slow down reconfiguration times in wide-area networks. Research has shown that micro-second switching of application traffic is possible in data center environments [54]–[56]. The ability to conduct circuit switching at microsecond timescales has illuminated further intrigue, particularly for transport protocols running on top of these networks. In C-Through [19], the authors observed that throughput for TCP applications dropped when their traffic migrated to the optical network. They showed how to mitigate this by increasing the queue size for optical circuit switches and adjusting the host behaviors. Mukerjee *et al.* [57] augmented their solution by expanding TCP for reconfigurable data center networks.

2) Algorithms: The capability of optical circuit switching for data center networks comes with the need to define new algorithms for optimizing utilization, bandwidth, fairness, latency, or any other metric of interest. Research has presented many different approaches for optimizing the metric relevant to the network operator in static networks. Traffic Engineering (TE) generally refers to the determination of paths for flows through the network, and the proportion of bandwidth levied for any particular flow. If the data center has a static network topology (*e.g.*, fat-tree), then TE is simple enough that switches can conclude how to route flows. However, introducing reconfigurable paths complicates the process of TE significantly: network elements (*e.g.*, switches) must now also determine with whom and when to establish optical paths, and when to change optical neighbors.

Matchings. Choosing where to establish optical circuits can be computed quickly via matching algorithms [58]. Matchings often provide a good approximation, especially in settings where the goal is to maximize single-hop throughput along with reconfigurable links. Matching algorithms hence frequently form the basis of reconfigurable optical networks, *e.g.*, *Helios* [59] and *c-Through* [19] rely on maximum matching algorithms. If there exist multiple reconfigurable links (say b many), it can be useful to directly work with a generalization of matching called b-matching [60]: b-matchings are for example used in *Proteus* [61] and its extension *OSA* [62]. In some scenarios, for example, when minimizing the average weighted path length under segregated routing, maximum b-matching algorithms even provide optimal results [63], [64]. This however is not always true, *e.g.*, when considering non-segregated routing policies [63], [64], which require heuristics [51, §5.1], [65].

Oblivious Approaches. Matchings also play a role in reconfigurable networks which *do not* account for the traffic they serve, *i.e.*, in *demand-oblivious* networks. The prime example here is *Rotornet* [18] which relies on a small set of matchings through which the network cycles endlessly: since these reconfigurations are "dumb", they are fast (compared to demand-aware networks) and provide frequent and periodic direct connections between nodes, which can significantly reduce infrastructure cost (also known as "bandwidth tax") compared to multihop routing. In case of uniform (delay-tolerant) traffic, such single-hop forwarding can saturate the network's bisection bandwidth [18]; for skewed traffic matrices, it can be useful to employ Valiant load balancing [66] to avoid underutilized direct connections, an idea recently also leveraged in *Sirius* [12] via Chang *et al.* [67]. *Opera* [68] extends *Rotornet* by maintaining expander graphs in its periodic reconfigurations. Even though the reconfiguration scheduling of *Opera* is deterministic and oblivious, the precomputation of the topology layouts in its current form is still randomized. Expander graphs (and their variants, such as random graphs [69]) are generally considered very powerful in data center contexts. An example of a demand-aware expander topology was proposed in *Tale of Two Topologies* [70], where the topology locally converts between Clos and random graphs.

Traffic Matrix Scheduling. Another general algorithmic approach is known as *traffic matrix scheduling*: the algorithmic optimizations are performed based on a snapshot of the demand, *i.e.*, based on a traffic matrix. For example, *Mordia* [71] is based on an algorithm that reconfigures the network multiple times for a single (traffic demand) snapshot. To this end, the traffic demand matrix is scaled into a bandwidth allocation matrix, which represents the fraction of bandwidth every possible matching edge should be allocated in an ideal schedule. Next, the allocation matrix is decomposed into a schedule, employing a computationally efficient [72] Birkhoff-von-Neumann decomposition, resulting in $O(n^2)$ reconfigurations and durations. This technique also applies to scheduling in hybrid data center networks which combine optical components with electrical ones, see *e.g.*, the heuristic used by *Solstice* [73]. *Eclipse* [74] uses traffic matrix scheduling to achieve a $(1 - 1/e^{(1-\varepsilon)})$-approximation for throughput in the hybrid switch architecture with reconfiguration delay, but only for direct routing along with single-hop reconfigurable connections. While *Eclipse* is an offline algorithm, Schwartz *et al.* [75] recently presented online greedy algorithms for this problem, achieving a provable competitive ratio over time; both algorithms allow to account for reconfiguration costs. Another example of traffic matrix scheduling is DANs [76]–[79] (short for demand-aware networks, which are optimized toward a given snapshot of the demand). DANs rely on concepts of demand-optimized data structures (such as biased binary search trees) and coding (such as Huffman coding) and typically aim to minimize the expected path length [76]–[79], or congestion [77]. In general, the problem features exciting connections to the scheduling literature, e.g., the work by Anand *et al.* [80], and more recently, Dinitz *et al.* [81]; the latter, however, is not based on matchings or bipartite graphs. Rater, the demands are the edges of a general graph, and a vertex cover can be communicated in each round. Each node can only send a certain number of packets in one round.

Self-Adjusting Datastructures. A potential drawback of traffic matrix scheduling algorithms is that without countermeasures, the optimal topology may change significantly from one traffic matrix snapshot to the next, even though the matrix is similar. There is a series of algorithms for reconfigurable networks which account for reconfiguration costs, by making a connection to self-adjusting data structures (such as splay trees) and coding (such as dynamic Huffman coding) [78], [82]–[85]. These networks react *quickly and locally* two new communication requests, aiming to strike an optimal trade-off between the benefits of reconfigurations (*e.g.*, shorter routes) and their costs (*e.g.*, reconfiguration latency, energy, packet reorderings, *etc.*).

To be more specific, the idea of both the traffic matrix scheduling algorithms and the self-adjusting data structure based algorithms is to organize the communication partners (*i.e.*, the destinations) of a *given* communication source in either a static binary search or Huffman tree (if the demand is known), or in a dynamic tree (if the demand is not known or if the distribution changes over time). The tree optimized for a single source is sometimes called the *ego-tree*, and the approach relies on combining these ego-trees of the different sources into a network while keeping the resulting node degree constant and preserving distances (*i.e.*, low distortion).

Machine Learning. Another natural approach to devise algorithms for reconfigurable optical networks is to use machine learning. To just give two examples, *xWeaver* [20] and *DeepConf* [86] use neural networks to provide traffic-driven topology adaptation. Another approach is taken by Kalmbach *et al.* [87], who aim to strike a balance between topology optimization and "keeping flexibilities", leveraging self-driving networks. Finally, Truong-Huu *et al.* [88] proposed an algorithm which uses a probabilistic, Markov-chain based model to rank ToR nodes in data centers as candidates for light-path creation.

Additional Aspects. Last but not least, several algorithms account for additional and practical aspects. In the context of shared mediums (*e.g.*, non-beamformed wireless broadcast, fiber[3] (rings)), contention and interference of signals can be avoided by using different channels and wavelengths. The algorithmic challenge is then to find (optimal) edge-colorings on multi-graphs, an NP-hard problem for which fast heuristics exist [90]. However, on specialized topologies, optimal solutions can be found in polynomial time, *e.g.*, in *WaveCube* [91]. Shared mediums also have the benefit that it is easier to distribute data in a one-to-many setting [92]. For example, on fiber rings, all nodes on the ring can intercept the signal [89, §3.1]. One-to-many paradigms[4] such as multicast can also be implemented in other technologies, using *e.g.*, optical splitters for optical circuit switches or half-reflection mirrors for free-space optics [95]–[99].

3) Systems Implementations: There have been many demonstrations of systems for reconfigurable optics in data centers. Many of the papers that we discuss in Section IV-B2 are fully operational systems. Another notable research development that does not fit into algorithms is the work by Mukerjee et al. [57]. They describe amendments to the TCP protocol to increase the efficiency of reconfigurable data center networks. These amendments include dynamic buffer re-sizing for switches and sharing explicit network feedback with hosts. Much of the other work on reconfigurable DCNs are summarized in Table II.

4) Open Challenges: Our understanding of algorithms and topologies in reconfigurable networks is still early, but first insights into efficient designs are being published. One front where much more research is required concerns the modeling (and dealing with) reconfiguration costs. Indeed, existing works differ significantly in their assumptions, even for the same technology, making it challenging to compare algorithms. Related to this is also the question of how reconfigurations affect other layers in the networking stack, and how to design (distributed) controllers. In terms of algorithms, even though a majority of problems are intractable to solve optimally, due to integral connection constraints, the question of approximation guarantees is mostly open. For example, consider designing a data center with minimum average weighted path length. A logarithmic approximation is easy to achieve by simply minimizing the diameter of a (constant-degree) static topology. However, computing an optimal solution is NP-hard. So, can we obtain polynomial approximation algorithms with

---

[3] In the context of data center proposals, shared fiber is the more popular medium, *e.g.*, in [62], [71], [89].

[4] Conceptually similar challenges arise for coflows [93], [94].

	Fabric	Dem.-Aw.	Novelty
Helios [59]	Hybrid	✓	First hybrid system using WDM for busty low-latency traffic
c-Through [19]	Hybrid	✓	Enlarged buffers for optical ports increases utilization
ProjecToR [52]	Hybrid/FSO	✓	Introduces DMDs for free-space switching thus enabling a fan-out potential to thousands of nodes
Proteus [61]	All-optical	✓	Design of an all-optical and reconfigurable DCN.
OSA [62]	All-optical	✓	Demonstrates greater reconfiguration flexibility and bisection bandwidth than hybrid architectures
Rotornet [18]	Hybrid	×	An all-optical demand-oblivious DCN architecture for simplified network management
Opera [68]	All-optical	×	Extends Rotornet to include expander graphs rotations
Flat-tree [70]	Hybrid	✓	A hybrid of random graphs and Clos topologies brings reconfigurable optics closer to existing DCNs.
Solstice [73]	Hybrid	✓	Exploits sparse traffic patterns in DCNs to achieve fast scheduling of reconfigurable networks.
Eclipse [100]	Hybrid	✓	Outperforms Solstice by applying sub-modular optimization theory to hybrid network scheduling.
xWeaver [20]	Hybrid	✓	Trains neural networks to construct performant topologies based on training data from historic traffic traces.
DeepConf [86]	Hybrid	✓	Presents a generic model for constructing learning systems of dynamic optical networks
WaveCube [91]	Hybrid	✓	A modular network architecture for supporting diverse traffic patterns.
Sirius [12]	All-optical	×	Achieves nanosecond-granularity reconfiguration for thousands of nodes

TABLE II: Summary of systems implementations of reconfigurable data center networks

constant performance trade-offs? Similarly, do good (fixed) parameter characterizations enable efficient run times, and what can we expect from e.g., linear time and distributed algorithms? Moreover, beyond general settings, how do specific (oblivious) network designs enable better algorithms, and how does their design interplay with topologies of the same equipment cost?

Next, going beyond scheduling, how can the framework of online algorithms be leveraged in this context? Ideally, we want a reconfigurable link to exist before *the traffic appears. How can we balance this from a worst-case perspective? In this context, traffic prediction techniques might reduce the possible solution space massively, but we will still need extremely rapid reaction times to new traffic information.*

Another open challenge is the efficient interplay between reconfigurable and non-reconfigurable network parts. Theory for specific reconfigurable topologies (e.g., traffic matrix scheduling for a single optical switch) has seen much progress. However, more general settings, particularly non-segregated routing onto both network parts, are still an open issue, beyond an abstract view of the combination with a single packet switch.

C. Reconfigurable Optical WAN

In this section, we survey recent research in reconfigurable optics in wide-area networks (WAN). Reconfigurable optics refers to dynamism in the physical-layer technology that enables high-speed and high-throughput WAN communications, fiber optics. We divide reconfigurable optical innovations into two sub-categories, rate-adaptive transceivers and dynamic optical paths. Rate adaptive transceivers are optical transceivers that can change their modulation format to adapt to physical layer impairments such as span-loss and noise. Dynamic optical paths refer to the ability to *steer light*, thus allowing the edges of the network graph to change (*e.g.*, to avoid a link that has failed).

Many groups have studied the programmability and autonomy of optical networks. Gringeri *et al.* [101] wrote a concise and illuminating introduction to the topic. In it, the authors propose extending Software Defined Network (SDN) principles to optical transport networks. They highlight challenges, such as reconfiguration latency in long-haul networks, and provide a trade-off characterization of distributed vs. centralized control for an optical SDN system. They claim that a tiered hierarchy of control for a multi-regional network (e.g., segregated optical and network control loops) will offer the best quality solution. Further, they argue that centralized control should work best to optimize competing demands across the network, but that the controller's latency will be too slow to react to network events, *e.g.*, link outages quickly. Therefore, the network devices should keep some functionality in their control plane to respond to link failures in a decentralized manner, *e.g.*, reallocating the lost wavelengths by negotiating an alternative path between the endpoints.

1) WAN-Specific Challenges and Solutions: There are many reasons for the prevalence of optical fiber as the de-facto leader for long-distance communications. First, it has incredible reach compared to copper—optical signals can propagate 80 to 100 km before being amplified. Second, it has an incredibly high bandwidth compared to the radio spectrum. Third, optical fiber itself has proved to be a robust medium over decades, as improvements to the transponders at the ends of the fiber have enabled operators to gain better value out of the same fiber year after year.

To design a WAN, the network architect must solve several difficult challenges, such as estimating the demand on the network now and into the future, optimal placement of routers and quantity of ports on those routers within the network, and optimal placement of amplifiers in the network.

Many design challenges solve more easily in a static WAN, where optical channels are initialized once and maintained for the network's life. For example, amplifiers carrying the channel must have their gain set in such a way that the signal is transmitted while maximizing the signal-to-noise ratio (SNR). This calculation can take minutes or hours depending on the network's characteristics (*e.g.*, the number of intermediate nodes and the number of distinct channels on shared amplifiers).

Dynamic optical networks must rapidly address these challenges (in sub-second time frames) to achieve the highest possible utilization, posing a significant challenge. For example, it requires multiple orders of magnitude increases in the provisioning time for optical circuits beyond what is typically offered by hardware vendors. Therefore, several research efforts have explored ways to automate WAN network elements' configuration concerning physical layer impairments in a robust and time-efficient manner.

Chromatic Dispersion. DWDM makes efficient use of optical fiber by putting as many distinct optical channels, each identified by a frequency (or lambda λ) onto the shared fiber. Each of these lambdas travels at a different speed relative to the speed of light. Therefore, two bits of information transmitted simultaneously via two different lambdas will arrive at the destination at two different times. Further, chromatic dispersion is also responsible for pulse-broadening, which reduces channel spacing between WDM channels and can cause FEC errors. Therefore, DWDM systems must handle this physical impairment.

Amplified Spontaneous Emission (ASE) Noise. A significant limitation of circuit switching is the latency of establishing the circuit due to ASE noise constraints [102]. Although SDN principles can apply to ROADMs and WSSs (to automate the control plane of these devices), physical layer properties, such as Noise Figure (NF) and Gain Flatness (GF) complicate the picture. When adding or removing optical channels to or from a long-haul span of fiber, traversing multiple amplifiers, the amplifiers on that path must adjust their gain settings to accommodate the new set of channels. To this end, researchers have worked to address the challenge of dynamically configuring amplifiers. Oliveira *et al.* [103] demonstrated how to control gain on EDFAs using GMPLS. They evaluated their solution on heterogeneous optical connections (10, 100, 200, and 400 Gbps) and modulations (OOK, QPSK, and 16-QAM). They used attenuators to disturb connections and allow their GMPLS control loop to adjust the amplifier's gains. They show that their control loop helps amplifiers to adjust while transmitting bits with BER below the FEC threshold for up to 6 dB of added attenuation.

Moura *et al.* [104] present a machine learning approach for configuring amplifier gain on optical circuits. Their approach uses case-based reasoning (CBR) as a foundation. The intuition behind CBR is that the gain setting for a set of circuits will be similar if similar circuits are present on a shared fiber. They present a genetic algorithm for configuring amplifiers based on their case-based reasoning assumption. They show that their methodology is suitable for configuring multiple amplifiers on a span with multiple optical channels. In a follow-up study, they present FAcCBR [105], an optimization of their genetic algorithm, which yields gain recommendations more quickly by limiting the number of data-points recorded by their algorithm.

Synchronization. Managing a WAN requires coordinating services (*e.g.*, end-to-end connections) among diverse sets of hardware appliances (transponders, amplifiers, routers), logically and consistently. The Internet Engineering Task Force (IETF) has defined protocols and standards for configuring WAN networks. As the needs and capabilities of networks have evolved, so have the protocols. Over the years, new protocols have been defined to bring more control and automation to the network operator's domain. These protocols are Simple Network Management Protocol (SNMP) [106] and Network Configuration Protocol (NETCONF) [107]. Additionally, network operators and hardware vendors have been working to define a set of generalized data models and configuration practices for automating WAN networks under the name OpenConfig [108]. Although OpenConfig is not currently standardized with the IETF, it is deployed and has demonstrated its value in several unique settings.

In addition to the standardized and proposed protocols for general-purpose WAN (re)configuration, there has been a push by various independent research groups to design and test protocols specifically for reserving and allocating optical channels in WAN networks.

One protocol was developed in conjunction with the CORONET [109] program, whose body of research has led to several other developments in reconfigurable optical WANs. The proposal, by *Skoog* et al. [110], describes a three-way handshake (3WHS) for reserving and establishing optical paths in single and multi-domain networks. In the 3WHS, messages are exchanged over an optical supervisory channel (OSC)—an out-of-band connection between devices isolated from user traffic. The transaction is initiated by one Optical Cross-Connect (OXC^A) and directed at a remote OXC, OXC^Z. At each hop along the way, the intermediate nodes append the available channels to the message. Then, OXC^Z chooses a channel via the first-fit strategy [111] and sends a message to OXC^A describing the chosen channel. Finally, OXC^A activates the chosen channel and beings sending data over it to OXC^Z. This protocol is claimed to meet the CORONET project standard for a setup time of 50 ms + RTT between nodes. Bit arrays are used to communicate the various potential channels between nodes and are processed in hardware. The blocking probability is 10^{-3} if there is one channel reserved between any two OXC elements so long as there are at least 28 total channels possible between OXCs [110].

2) Algorithms: Jointly optimizing both the optical and the network layer in wide-area networks leads to new opportunities to improve performance and efficiency, while introducing new algorithmic challenges. In contrast to the previously discussed data center networks, it is impossible to create new topological connections in a wide-area network (without deploying more fiber. Free space optics solutions don't apply here). Instead, reconfigurability is possible by adjusting and shifting bandwidth capacities along the fiber edges, possibly over multiple hops. Hence, we need a different set of algorithmic ideas that optimize standard metrics such as throughput, completion time, blocking probability, and resilience. In this section, we discuss recent papers that tackle these issues, starting with some earlier ones. Moreover, there is the need for some central control to apply the routing, policy, lightpath *etc.* changes, for which we refer to recent surveys [112], [113].

Routing aspects are explored intensively in this context. Algorithmic approaches to managing reconfigurable optical topologies have been studied for decade, but are recently gaining new attention. In early work by Kodialam *et al.* [114] explores IP and optical wavelength routing for a series of connection requests. Their algorithm determines whether a request should be routed over the existing IP topology, or if a new optical path should be provisioned for it. Interesting work by Brzezinski and Modiano [115] who leverage matching algorithms and Birkhoff–von Neumann matrix decompositions and evaluate multi- versus single-hop routing[5] in WDM networks under stochastic traffic. However, the authors mostly consider relatively small networks, *e.g.*, with three to six nodes. For larger networks, shortest lightpath routing is a popular choice [117]. Another fundamental aspect frequently considered in the literature regards resilience [118]–[120]. For example, Xu *et al.* [118] investigate resilience in the context of shared risk link groups (SLRGs) and propose a method on how to provision the circuits in a WAN. To this end, they construct Integer Linear Programs to obtain maximally SLRG-diverse routes, which they then augment with post-processing for DWDM system selection and network design issues.

We now introduce further selected algorithmic works, starting with the topic of bulk transfers [121]. In *OWAN* [122], Jin *et al.* optimize bulk transfers in a cross-layer approach, which leverages both the optical and the network layer. Their main objective is to improve completion time; while an integer linear program formulation would be too slow, the authors rely on a simulated annealing approach. A local search shifts wavelength allocations, allowing heuristic improvements to be computed at a sub-second scale. The scheduling of the bulk transfer then follows the standard shortest job first approaches. When updating the network state, if desired, *OWAN* can extend prior consistent network update solutions [123] by introducing circuit nodes in the corresponding dependency graphs. *OWAN* also considers deadline constrained traffic, implementing the earliest deadline first policy. Follow-up work extended *OWAN* in two directions, via theoretic scheduling results and for improvements on deadline-constrained transfers.

In *DaRTree* [124], Luo *et al.* develop an appropriate relaxation of the cross-layer optimization problem for bulk transfers under deadlines. Their approach relies on a non-greedy allocation in an online setting, which allows future transfers to be scheduled efficiently without needing to reallocate currently utilized wavelengths. To enhance multicast transfers (*e.g.*, for replication), they develop load-adaptive Steiner Tree heuristics.

Jia *et al.* [125] design various online scheduling algorithms and prove their competitiveness in the setting of *OWAN* [122]. The authors consider the minimum makespan and sum completion time, analyzing and extending greedy cross-layer scheduling algorithms, achieving small competitive ratios. Dinitz and Moseley [126] extend the work of Jia *et al.* by considering a different objective, the sum of flow times in an online setting. They show that resource augmentation is necessary for acceptable competitive bounds in this setting, leading to nearly (offline) optimal competitive ratios. While their algorithms are easy to implement (*e.g.*, relying on ordering by release time or by job density), the analysis is complicated and relies on linear program relaxations. Moreover, their algorithm also allows for constant approximations in the weighted completion time setting, without augmentations.

Another (algorithmic) challenge is the integration of cross-layer algorithms into current traffic engineering systems. Such TEs are tried and tested, and hence service providers are reluctant to adapt their designs. To this end, Singh *et al.* [127] propose an abstraction on how dynamic link capacities (*e.g.*, via bandwidth variable transceivers) can be inserted into classic TEs. Even though the TE is oblivious to the optical layer, an augmentation of the IP layer with fake links enables cross-layer optimization via the TE. A proposal [128] for a new TE for such dynamic link capacities is discussed in the next Section IV-C3. Singh *et al.* [127] also discuss consistent update methods [129] for dynamic link capacities, which Tseng [130] formalizes into a rate adaption planning problem, providing intractability results and an LP-based heuristic.

OptFlow [131] proposes a cross-layer abstraction for programmable topologies as well, but focuses on shifting wavelengths between neighboring fibers. Here, the abstraction concept is extended by not only creating fake links but also augmenting the traffic matrix with additional flows. As both links and flows are part of the input for TEs, *OptFlow* enables the compilation of optical components into the IP layer for various traffic engineering objectives and constraints. Concerning consistent updates, classic flow-based techniques [129] carry over, enabling consistent cross-layer network updates too.

Higher-layer applications, such as VNF Network Embedding (VNF-NE) have also been studied by various groups [132], [133]. Network embedding is a physical layer abstraction for creating end-to-end paths for network applications or network function virtualization (NFV) service chains. Paths have requirements for both bandwidth and CPU resources along the service chain. Wang *et al.* [132] proves this problem to be NP-Complete for elastic optical networks. Soto *et al.* [133] provides an integer linear program (ILP) to solve the VNF-NE problem. The ILP solution is intractable for large networks. Thus they provide a heuristic that uses a ranking-system for optical paths. Their heuristic ranks optical paths by considering a set of end-to-end connection requests. Paths with higher rank satisfy a more significant proportion of the demand for bandwidth and CPU among all of the requests.

Optical layer routing with traffic and application constraints is a difficult problem. The running theme has been that linear programming solutions can find provably optimal solutions [134], but take too long to converge for most use cases.

---

[5] See also the idea of lightpath splitting in Elastic Optical Networks [116].

However, network traffic is not entirely random and therefore has an underlying structure that may be exploited by offline linear program solvers, as shown by Kokkinos *et al.* [135]. They use a two-stage approach for routing optical paths in an online manner. Their technique finds periodic patterns over an epoch (e.g., daily, weekly, or monthly) and solves the demand characterized within the epoch with an offline linear program. Then, their online heuristic makes changes to the topology to accommodate random changes in demand within the epoch.

3) Systems Implementations: The integration of reconfigurable optics with WAN systems has been impracticable due to its cost and a lack of convergence on cross-layer APIs for managing the WAN optical layer with popular SDN controllers. However, some exciting work has demonstrated the promise for reconfigurable optics is closed settings. Notably, RADWAN [136] and CORONET [109] for bandwidth-variable WAN systems and systems with dynamic optical paths, respectively. In this section, we explore reconfigurable optical WAN systems more deeply in these two contexts. Table III summarizes these systems.

Bandwidth Variable Transceivers. A team of researchers at Microsoft evaluates bandwidth variable transponders' applicability for increases throughput in Azure's backbone in North America [46]. They find that throughput for the WAN can increase if they replace the fixed-rate transponders in their backbone network with three-way sliceable transponders. They also show that for higher-order slices, bandwidth gran increases at diminishing returns.

Traffic Engineering with rate-adaptive transceivers was recently proposed by Singh *et al.* [128]. The authors are motivated by a data-set of Microsoft's WAN backbone Signal-to-Noise ratio from all transceivers in the North-American backbone, over two and a half years. They note that over 60% of links in the network could operate at 0.75× higher capacity and that 25% of observed outages due to SNR drops could be mitigated by reducing the modulation of the affected transceivers. They evaluate the reconfigurability of Bandwidth-Variable Transponders, showing that reconfiguration time for the transceivers could be reduced from minutes to milliseconds by *not* turning the transceivers off. Then, they propose a TE objective function via linear-programming, to minimize churn, or impact due to SNR fluctuations, in a WAN. Finally, they evaluate their TE controller on a testbed WAN and show that they improve network throughput by 40% over a competitive software-defined networking controller, SWAN [138].

Dynamic Optical Paths. In the early aughts, researchers explored the benefit of dynamic optical paths for networks in the context of *grid-computing*. Early efforts by Figueira *et al.* [139] addressed how a system might manage dynamic optical paths in networks. In this work, the authors propose a web-based interface for submitting optical reconfiguration requests and a controller for optimizing the requests' fulfillment. They evaluate their system on OMNInet [140], a metropolitan area network with 10 Gbps interconnects between 4 nodes and Wavelength Selective Switches between them. They claim that they can construct optical circuits between the OMNInet nodes in 48 seconds. Further, they show that amortized setup time and transfer is faster than packet-switching for files 2.5 Gb or larger (assuming 1 Gbps or greater optical interconnect and 300 Mbps packet switching throughput). They go on to evaluate file transfer speeds using the optical interconnect and show that they can archive average transfer speeds of 680 Gbps. Iovanna *et al.* [141] address practical aspects of managing multilayer packet-optical systems. They present a set of useful abstractions for operating reconfigurable optical paths in traffic engineering using an existing management protocol, GMPLS.

Stability is an important feature of any network. An interesting question about reconfigurable optical networked systems arises regarding the stability of optically switched paths. That is, if the topology can continuously change to accommodate random requests, what service guarantees can the network make? Can the fluctuation of the optical layer be detrimental to IP layer services? Chamania *et al.* [142] explore this issue in detail, providing an optimal solution for keep quality of service guarantees for IP traffic while also improving performance beyond static optical layer systems.

Bandwidth-on-demand (BoD) is an exciting application of reconfigurable networks. Von Lehmen *et al.* [109] describe their experience in deploying BoD services on CORONET, DARPA's WAN backbone. They implement protocols for add/dropping wavelengths in their WAN with a novel 3-way-handshake protocol. They demonstrate how their system is able to utilize SWAN [138] Traffic Engineering Controller as one such application that benefits from the BoD service.

More recently, there has been a resurgence academic work highlighting the potential benefit of dynamic optical paths in the WAN. One such system, called OWAN (Optical Wide-Area Network) [143], proposes how to use dynamic optical paths to improve the delivery time for bulk transfers between data centers. They build a testbed network with home-built ROADMs and implement a TE controller to orchestrate bulk transfers between hosts in a mesh optical network of nine nodes. They compare their results with other state-of-the-art TE systems, emphasizing that OWAN delivers more transfers *on time* than any other competing methods.

Dynamic optical paths increase the complexity of networks and capacity planning tasks because any optical fiber may need to accommodate diverse and variable channels. However, this complexity is rewarded with robustness or tolerance to fiber link outages. Gossels *et al.* [144] propose dynamic optical paths to make long-haul networks more robust and resilient to node and link failures by presenting algorithms for allocating bandwidth on optical paths dynamically in a mesh network. Their objective is to protect networks from any single node or link failure event. To this end, they present an optimization framework for network planners, which determines where to deploy transponders to minimize costs while running a network over dynamic optical paths.

Another effort in reducing the complexity of dynamic optical path WAN systems was presented by Dukic *et al.* [11]. Their system, *Iris*, exploits a unique property of regional connectivity, *i.e.*, the vast abundance of optical fiber in dense metropolitan areas [145]. They find that the complexity of managing dynamic optical paths is greatly reduced when switching at the fiber-strand level versus the (sub-fiber) wave-

	BVT	Network Design	Amps.	Algorithms
CORONET [109]	×	×	×	3WHS
OWAN [122]	×	×	×	Simulated Annealing
FACcBR [104]	×	×	✓	Case Based Reasoning
RADWAN [136]	✓	×	×	Linear Program
DDN [137]	×	✓	✓	Time-slotted packet scheduling
Iris [11]	×	✓	✓	Shortest path for any failure scenario

TABLE III: Summary of systems implementations of reconfigurable wide area networks

length level. To this end, they detail their design trade-off space for inter-data center connectivity across metropolitan areas. They deploy their system in a hardware testbed to emulate connectivity between three data centers, verifying that optical switching can be done in 50 to 70 ms over three amplifiers. They obviate amplifier reconfiguration delays by conducting fiber-level switching rather than wavelength-level. Thus, the amplifiers on a fiber path are configured once for the channel that traverses it. When a circuit changes its path, away from one data center and towards another, it uses a series of amplifiers that have been pre-configured to accommodate the loss of that given circuit.

Inter-data center network connectivity over a regional optical backbone was also investigated by Benzaoui *et al.* [137]. Their system, *Deterministic Dynamic Network (DDN)*, imposes strict constraints for application layer latency and jitter. They show that they can reconfigure optical links in under 2 ms, and guarantee consistent latency and jitter through their time-slotted scheduling approach.

4) Open Challenges: Wide-area optical networks are rich with open challenges. The works presented in this section highlight significant developments that have been made towards reconfigurable WAN systems and illuminate great benefits for such systems. However, programmability, cross-layer information sharing, and physical impairment problems must be solved. On the programmability front, efforts such as OpenConfig [108], OpenROADM [146], and ONOS [147] are working to provide white-box system stacks for optical layer equipment. If these are widely adopted and standardized, this will open the door for agile and efficient use of wide-area networks for a variety of applications (e.g., new tools to combat DDoS [148]). Other challenges include wrangling with the physical constraints of efficient and rapidly reconfigurable WANs, for example, coordination of power adjustments across amplifiers for long-haul circuits in conjunction with traffic engineering workloads and constraints.

V. FUTURE WORK AND APPLICATIONS

Having looked into the theory and practice, my thesis to demonstrate the benefits of cross-layer programmability along four dimensions: network simulation, network measurement, traffic engineering, and cyber security. The following subsections each highlight a potential future application in each category.

A. Vertically Programmable Network Simulator

Recall the challenge from section III, that there is no clear simulation framework for evaluating and constructing reconfigurable optical networks. The lack of such a framework makes it difficult to compare systems with and without dynamic optical components. My future work is to address this challenge by developing a vertically programmable network simulator. The purpose of this simulator is to provide a means to compare optical layer reconfiguration strategies in conjunction with higher layer packet-switching and routing strategies. I hope to publish this in a simulation venue (*e.g.*, MASCOTS or SummerSim).

B. Network Measurement

In section IV-A, one of the main challenges was the transparency of the optical-layer network. This transparency has two adverse effects. First, it makes active-measurement based mapping of the physical network impossible (the WSSes and ROADMs are invisible to network measurement). Second, our fiber optical network's control and management functions are rendered more complex by lack of a complete mapping between the physical and logical (router-level) topology. My future work is to address this challenge by proposing a set of secure enclaves for monitoring optical networks from the network layer. This hypothetical system is an *optical-layer traceroute* or OLT. OLT's main idea is to make the optical layer visible to higher-layer applications (e.g., traffic engineering or TE). The critical requirements for secure enclaves are to make the optical layer visible while also preserving the network's privacy. We can achieve this through a set of *software hooks* that expose information about optical line equipment. Types of information that these hooks could safely include are unique hashes of object IDs and optical layer performance information (e.g., bit error rate or BER). I hope to publish this work at a network measurement venue (*e.g.*, IMC or PAM).

C. Traffic Engineering (TE)

Recall the challenges from section IV-C. First, we must overcome the significant performance penalties for reconfiguration imposed by optical equipment (such as amplifiers and transponders) deployed in today's WANs. Second, our ability to dynamically reconfigure optical devices must be utilized rapidly by *optical path protection schemes* (e.g., amplifier gain control or AGC, transponder power adjustments) and by TE algorithms that operate above the physical layer. Third, there is no unified formulation to assess reconfigurable optical network paths vs. static allocation or quantify how existing TE schemes can benefit from static backup paths for optimally routing flows through the network.

I'm addressing these challenges in my ongoing and future work. Regarding the significant performance penalties for long-haul link changes, I am conducting a study on amplifier reconfiguration times to determine bottlenecks in the process and find methods for removing those bottlenecks. I'm going to publish this study at the Optical Fiber Conference (OFC). Regarding the timely activation of optical paths for applications such as path-protection schemes, we require a vertically programmable network control system to view and manage several aspects of the network (*e.g.*, active wavelength connections, traffic demand on the wavelengths, and the physical topology components and resources). I'm building a simulator to accurately model the system's inputs, the dynamics for optical layer reconfiguration, and its effect on traffic and concrete objective functions. This work overlaps with the goals for a vertically programmable network simulator from V-A. I hope to publish this at a systems venue (*e.g.*, NSDI or SIGCOMM).

D. Cybersecurity

We are now in the era of terabit DDoS attacks. The immense attack volumes, attack diversity, sophisticated attack strategies, and the low cost of launching them make DDoS attacks a critical cybersecurity issue in Internet infrastructure now and for the foreseeable future. DDoS is not a new problem, and prior work has made significant progress in devising mitigation strategies to tackle DDoS attacks. These include packet scrubbing solutions, in-network filtering, and more recent routing around congestion methods [149], as well SDN/NFV-based elastic defenses [150], [151]. Despite these advances, the rise of new-age extreme terabit attacks mandates a critical rethinking of DDoS defense strategies.

In my current and future work, I am exploring a new opportunity for bolstering our DDoS defense arsenal by leveraging advances in programmable optics, which I highlighted in this area exam. The main idea is that our ability to program a subset of the optical components provides an untapped defense resource. I'm working to show how topology changes can improve current defense solutions such as hardware-based packet scrubbing and SDN/NFV defense. I hope to publish this at USENIX Security.

VI. CONCLUSION AND GENERAL OPEN CHALLENGES

Reconfigurable optical networks are a young technology, and much of their potential and limitations are not well-understood today. A first open challenge regards the identification of the killer use cases of reconfigurable technologies. In this paper, we have specifically considered data center and wide-area networks. Still, many other networks may benefit from similar technologies, and even in our context, the tradeoffs between costs and benefits (*e.g.*, in terms of resilience, performance, efficiency) are not well understood. In particular, these tradeoffs also depend on the specific technology, *e.g.*, on the reconfiguration time, as well as the traffic pattern; for example, demand-aware reconfigurable networks may only be useful if the traffic pattern exhibits temporal and spatial structure [152]. We currently specifically lack models for reconfiguration costs, and these costs, in turn, depend on the control plane, which is another open research challenge. It is not clear whether decentralized control planes are always superior to centralized ones, or whether hybrid designs are required. It is also not clear how to optimally design such control planes. From an algorithmic point of view, reconfigurable optical networks present a mostly uncharted complexity landscape. Whereas classic networking problems can largely rely on decades of optimization and graph theory, reconfiguration adds new and different twists to networking problems. In addition to surveying the research in reconfigurable optical networks, we have also identified exciting areas of intrigue, both for established areas and for novel applications.

REFERENCES

[1] S. Jain, A. Kumar, S. Mandal, J. Ong, L. Poutievski, A. Singh, S. Venkata, J. Wanderer, J. Zhou, M. Zhu *et al.*, "B4: Experience with a Globally-deployed Software Defined WAN," *ACM SIGCOMM*, 2013.

[2] C.-Y. Hong, S. Kandula, R. Mahajan, M. Zhang, V. Gill, M. Nanduri, and R. Wattenhofer, "Achieving high utilization with software-driven wan," in *SIGCOMM*, vol. 43, no. 4, 2013, pp. 15–26.

[3] P. Patel, D. Bansal, L. Yuan, A. Murthy, A. Greenberg, D. A. Maltz, R. Kern, H. Kumar, M. Zikos, H. Wu, C. Kim, and N. Karri, "Ananta: Cloud scale load balancing," in *SIGCOMM*, 2013. [Online]. Available: http://doi.acm.org/10.1145/2486001.2486026

[4] D. E. Eisenbud, C. Yi, C. Contavalli, C. Smith, R. Kononov, E. Mann-Hielscher, A. Cilingiroglu, B. Cheyney, W. Shang, and J. D. Hosein, "Maglev: A fast and reliable software network load balancer," in *NSDI*, 2016. [Online]. Available: https://www.usenix.org/conference/nsdi16/technical-sessions/presentation/eisenbud

[5] P. Bosshart, D. Daly, G. Gibb, M. Izzard, N. McKeown, J. Rexford, C. Schlesinger, D. Talayco, A. Vahdat, G. Varghese, and D. Walker, "P4: Programming Protocol-independent Packet Processors," *SIGCOMM*, vol. 44, no. 3, pp. 87–95, Jul. 2014. [Online]. Available: http://doi.acm.org/10.1145/2656877.2656890

[6] N. McKeown, T. Anderson, H. Balakrishnan, G. Parulkar, L. Peterson, J. Rexford, S. Shenker, and J. Turner, "Openflow: enabling innovation in campus networks," *SIGCOMM*, 2008.

[7] D. Firestone, A. Putnam, S. Mundkur, D. Chiou, A. Dabagh, M. Andrewartha, H. Angepat, V. Bhanu, A. Caulfield, E. Chung *et al.*, "Azure accelerated networking: Smartnics in the public cloud," in *NSDI*, 2018, pp. 51–66.

[8] A. S. Thyagaturu, A. Mercian, M. P. McGarry, M. Reisslein, and W. Kellerer, "Software defined optical networks (sdons): A comprehensive survey," *IEEE Communications Surveys & Tutorials*, 2016.

[9] B. C. Chatterjee, N. Sarma, and E. Oki, "Routing and spectrum allocation in elastic optical networks: A tutorial," *IEEE Communications Surveys Tutorials*, vol. 17, no. 3, pp. 1776–1800, 2015.

[10] D. M. Marom, P. D. Colbourne, A. D'errico, N. K. Fontaine, Y. Ikuma, R. Proietti, L. Zong, J. M. Rivas-Moscoso, and I. Tomkos, "Survey of photonic switching architectures and technologies in support of spatially and spectrally flexible optical networking," *IEEE/OSA Journal of Optical Communications and Networking*, vol. 9, no. 1, pp. 1–26, 2017.

[11] V. Dukic, G. Khanna, C. Gkantsidis, T. Karagiannis, F. Parmigiani, A. Singla, M. Filer, J. L. Cox, A. Ptasznik, N. Harland, W. Saunders, and C. Belady, "Beyond the mega-data center: networking multi-data center regions," in *SIGCOMM*. ACM, 2020, pp. 765–781.

[12] H. Ballani, P. Costa, R. Behrendt, D. Cletheroe, I. Haller, K. Jozwik, F. Karinou, S. Lange, K. Shi, B. Thomsen, and H. Williams, "Sirius: A flat datacenter network with nanosecond optical switching," in *SIGCOMM*. ACM, 2020, pp. 782–797.

[13] K.-T. Foerster and S. Schmid, "Survey of reconfigurable data center networks: Enablers, algorithms, complexity," *SIGACT News*, vol. 50, no. 2, pp. 62–79, Jul. 2019. [Online]. Available: http://doi.acm.org/10.1145/3351452.3351464

[14] A. Celik, B. Shihada, and M.-S. Alouini, "Optical wireless data center networks: potentials, limitations, and prospects," in *Broadband Access Communication Technologies XIII*, vol. 10945. International Society for Optics and Photonics, 2019, p. 109450I.

[15] Y. Lu and H. Gu, "Flexible and scalable optical interconnects for data centers: Trends and challenges," *IEEE Communications Magazine*, vol. 57, no. 10, pp. 27–33, October 2019.

[16] A. Willner, *Optical fiber telecommunications*. Academic Press, 2019, vol. 11.

[17] J. A. Bannister, L. Fratta, and M. Gerla, "Topological design of the wavelength-division optical network," in *IEEE INFOCOM'90*. IEEE Computer Society, 1990, pp. 1005–1006.

[18] W. M. Mellette, R. McGuinness, A. Roy, A. Forencich, G. Papen, A. C. Snoeren, and G. Porter, "RotorNet: A Scalable, Low-complexity, Optical Datacenter Network," in *SIGCOMM*, 2017.

[19] G. Wang, D. G. Andersen, M. Kaminsky, K. Papagiannaki, T. E. Ng, M. Kozuch, and M. Ryan, "c-through: Part-time optics in data centers," in *SIGCOMM*, 2010, pp. 327–338.

[20] M. Wang, Y. Cui, S. Xiao, X. Wang, D. Yang, K. Chen, and J. Zhu, "Neural network meets DCN: traffic-driven topology adaptation with deep learning," *POMACS*, vol. 2, no. 2, pp. 26:1–26:25, 2018.

[21] S. Lange, A. Raja, K. Shi, M. Karpov, R. Behrendt, D. Cletheroe, I. Haller, F. Karinou, X. Fu, J. Liu, A. . Lukashchuk, B. Thomsen, K. Jozwik, P. Costa, T. J. Kippenberg, and H. Ballani, "Sub-nanosecond optical switching using chip-based soliton microcombs," in *Optical Fiber Communication Conference (OFC'20)*. The Optical Society (OSA), March 2020.

[22] M. Ghobadi, R. Mahajan, A. Phanishayee, P.-A. Blanche, H. Rastegarfar, M. Glick, and D. Kilper, "Design of mirror assembly for an agile reconfigurable data center interconnect," Tech. Rep. MSR-TR-2016-33, 2016.

[23] A. Shakeri, M. Garrich, A. Bravalheri, D. Careglio, J. Solé-Pareta, and A. Fumagalli, "Traffic allocation strategies in wss-based dynamic optical networks," *Journal of Optical Communications and Networking*, vol. 9, no. 4, pp. B112–B123, 2017.

[24] G. Wang, D. G. Andersen, M. Kaminsky, M. Kozuch, T. S. E. Ng, K. Papagiannaki, M. Glick, and L. B. Mummert, "Your data center is a router: The case for reconfigurable optical circuit switched paths," in *HotNets*. ACM SIGCOMM, 2009.

[25] R. Durairajan, P. Barford, J. Sommers, and W. Willinger, "Intertubes: A study of the US long-haul fiber-optic infrastructure," in *SIGCOMM*. ACM, 2015, pp. 565–578.

[26] C.-Y. Hong, S. Mandal, M. Al-Fares, M. Zhu, R. Alimi, K. N. B., C. Bhagat, S. Jain, J. Kaimal, S. Liang, K. Mendelev, S. Padgett, F. Rabe, S. Ray, M. Tewari, M. Tierney, M. Zahn, J. Zolla, J. Ong, and A. Vahdat, "B4 and after: Managing hierarchy, partitioning, and asymmetry for availability and scale in google's software-defined wan," in *the 2018 Conference of the ACM Special Interest Group on Data Communication*, ser. SIGCOMM '18. New York, NY, USA: Association for Computing Machinery, 2018, p. 74–87.

[27] I. Tsirilakis, C. Mas, and I. Tomkos, "Cost comparison of ip/wdm vs. ip/otn for european backbone networks," in *2005 7th International Conference Transparent Optical Networks, 2005.*, vol. 2. IEEE, 2005, pp. 46–49.

[28] O. Gerstel, C. Filsfils, T. Telkamp, M. Gunkel, M. Horneffer, V. Lopez, and A. Mayoral, "Multi-layer capacity planning for ip-optical networks," *IEEE Communications Magazine*, vol. 52, no. 1, pp. 44–51, 2014.

[29] M. Gunkel, A. Autenrieth, M. Neugirg, and J.-P. Elbers, "Advanced multilayer resilience scheme with optical restoration for ip-over-dwdm core networks," in *2012 IV International Congress on Ultra Modern Telecommunications and Control Systems*. IEEE, 2012, pp. 657–662.

[30] P. Papanikolaou, K. Christodoulopoulos, and E. Varvarigos, "Joint multilayer planning of survivable elastic optical networks," in *Optical Fiber Communication Conference*. Optical Society of America, 2016, pp. M2K–3.

[31] S. Kozdrowski, M. Żotkiewicz, and S. Sujecki, "Optimization of optical networks based on cdc-roadm technology," *Applied Sciences*, vol. 9, no. 3, p. 399, 2019.

[32] O. Turkcu and S. Subramaniam, "Performance of optical networks with limited reconfigurability," *IEEE/ACM Trans. Netw.*, vol. 17, no. 6, pp. 2002–2013, Dec. 2009. [Online]. Available: http://dx.doi.org/10.1109/TNET.2009.2014158

[33] S. L. Woodward, M. D. Feuer, I. Kim, P. Palacharla, X. Wang, and D. Bihon, "Service velocity: Rapid provisioning strategies in optical roadm networks," *Journal of Optical Communications and Networking*, vol. 4, no. 2, pp. 92–98, 2012.

[34] X. Chen, Z. Zhu, R. Proietti, and S. J. B. Yoo, "On incentive-driven vnf service chaining in inter-datacenter elastic optical networks: A hierarchical game-theoretic mechanism," *IEEE Transactions on Network and Service Management*, vol. 16, no. 1, pp. 1–12, 2019.

[35] X. J. Zhang, S. Kim, and S. S. Lumetta, "Opportunity cost analysis for dynamic wavelength routed mesh networks," *IEEE/ACM Transactions on Networking*, vol. 19, no. 3, pp. 747–759, 2011.

[36] M. Team, "Mininet: An Instant Virtual Network on Your Laptop (or Other PC)," 2012.

[37] P. Kumar, C. Yu, Y. Yuan, N. Foster, R. Kleinberg, and R. Soulé, "Yates: Rapid prototyping for traffic engineering systems," in *Proceedings of the Symposium on SDN Research*, 2018, pp. 1–7.

[38] C. Kachris and I. Tomkos, "A survey on optical interconnects for data centers," *IEEE Communications Surveys & Tutorials*, vol. 14, no. 4, pp. 1021–1036, 2012.

[39] J. M. Taboada, J. J. Maki, S. Tang, L. Sun, D. An, X. Lu, and R. T. Chen, "Thermo-optically tuned cascaded polymer waveguide taps," *Applied physics letters*, vol. 75, no. 2, pp. 163–165, 1999.

[40] A. Yeniay, R. Gao, K. Takayama, R. Gao, and A. F. Garito, "Ultra-low-loss polymer waveguides," *Journal of lightwave technology*, vol. 22, no. 1, pp. 154–158, 2004.

[41] D. De Felipe, M. Kleinert, C. Zawadzki, A. Polatynski, G. Irmscher, W. Brinker, M. Moehrle, H.-G. Bach, N. Keil, and M. Schell, "Recent developments in polymer-based photonic components for disruptive capacity upgrade in data centers," *Journal of Lightwave Technology*, vol. 35, no. 4, pp. 683–689, 2016.

[42] K. A. Clark, D. Cletheroe, T. Gerard, I. Haller, K. Jozwik, K. Shi, B. Thomsen, H. Williams, G. Zervas, H. Ballani *et al.*, "Synchronous subnanosecond clock and data recovery for optically switched data centres using clock phase caching," *Nature Electronics*, vol. 3, no. 7, pp. 426–433, 2020.

[43] H. Toshiyoshi and H. Fujita, "Electrostatic micro torsion mirrors for an optical switch matrix," *Journal of Microelectromechanical systems*, vol. 5, no. 4, pp. 231–237, 1996.

[44] N. Farrington, G. Porter, S. Radhakrishnan, H. H. Bazzaz, V. Subramanya, Y. Fainman, G. Papen, and A. Vahdat, "Helios: a hybrid electrical/optical switch architecture for modular data centers," in *SIGCOMM*, 2010, pp. 339–350.

[45] G. Baxter, S. Frisken, D. Abakoumov, H. Zhou, I. Clarke, A. Bartos, and S. Poole, "Highly programmable wavelength selective switch based on liquid crystal on silicon switching elements," in *2006 Optical Fiber Communication Conference and the National Fiber Optic Engineers Conference*. IEEE, 2006, pp. 3–pp.

[46] M. Filer, J. Gaudette, M. Ghobadi, R. Mahajan, T. Issenhuth, B. Klinkers, and J. Cox, "Elastic optical networking in the microsoft cloud," *IEEE/OSA Journal of Optical Communications and Networking*, vol. 8, no. 7, pp. A45–A54, 2016.

[47] M. Jinno, B. Kozicki, H. Takara, A. Watanabe, Y. Sone, T. Tanaka, and A. Hirano, "Distance-adaptive spectrum resource allocation in spectrum-sliced elastic optical path network [topics in optical communications]," *IEEE Communications Magazine*, vol. 48, no. 8, pp. 138–145, August 2010.

[48] N. Sambo, P. Castoldi, A. D'Errico, E. Riccardi, A. Pagano, M. S. Moreolo, J. M. Fabrega, D. Rafique, A. Napoli, S. Frigerio *et al.*, "Next generation sliceable bandwidth variable transponders," *IEEE Communications Magazine*, vol. 53, no. 2, pp. 163–171, 2015.

[49] D. Thomson, A. Zilkie, J. E. Bowers, T. Komljenovic, G. T. Reed, L. Vivien, D. Marris-Morini, E. Cassan, L. Virot, J.-M. Fédéli *et al.*, "Roadmap on silicon photonics," *Journal of Optics*, vol. 18, no. 7, p. 073003, 2016.

[50] H. Li, H.-y. Zhang, L. Wang, Y.-b. Li, J.-s. Lai, R. Tang, W.-y. Zhao, B.-b. Wu, D. Wang, X. Zhao *et al.*, "Field trial of network survivability based on otn and roadm hybrid networking," in *Asia Communications and Photonics Conference*. Optical Society of America, 2017, pp. M3C–2.

[51] N. H. Azimi, Z. A. Qazi, H. Gupta, V. Sekar, S. R. Das, J. P. Longtin, H. Shah, and A. Tanwer, "Firefly: a reconfigurable wireless data center fabric using free-space optics," in *SIGCOMM*. ACM, 2014, pp. 319–330.

[52] M. Ghobadi, R. Mahajan, A. Phanishayee, N. R. Devanur, J. Kulkarni, G. Ranade, P. Blanche, H. Rastegarfar, M. Glick, and D. C. Kilper, "Projector: Agile reconfigurable data center interconnect," in *SIGCOMM*. ACM, 2016.

[53] S. Bloom, E. Korevaar, J. Schuster, and H. Willebrand, "Understanding the performance of free-space optics," *Journal of optical Networking*, vol. 2, no. 6, pp. 178–200, 2003.

[54] N. Farrington, G. Porter, Y. Fainman, G. Papen, and A. Vahdat, "Hunting mice with microsecond circuit switches," in *HotNets*. ACM, 2012, pp. 115–120.

[55] G. Porter, R. Strong, N. Farrington, A. Forencich, P. Chen-Sun, T. Rosing, Y. Fainman, G. Papen, and A. Vahdat, "Integrating microsecond circuit switching into the data center," in *SIGCOMM'13*, 2013, pp. 447–458.

[56] N. Farrington, A. Forencich, G. Porter, P. . Sun, J. E. Ford, Y. Fainman, G. C. Papen, and A. Vahdat, "A multiport microsecond optical circuit switch for data center networking," *IEEE Photonics Technology Letters*, vol. 25, no. 16, pp. 1589–1592, Aug 2013.

[57] M. K. Mukerjee, C. Canel, W. Wang, D. Kim, S. Seshan, and A. C. Snoeren, "Adapting tcp for reconfigurable datacenter networks," in *NSDI*, 2020, pp. 651–666.

[58] J. Edmonds, "Paths, trees and flowers," *Canad. J. Math*, vol. 17, pp. 449–467, 1965.

[59] N. Farrington, G. Porter, S. Radhakrishnan, H. H. Bazzaz, V. Subramanya, Y. Fainman, G. Papen, and A. Vahdat, "Helios: a hybrid electrical/optical switch architecture for modular data centers," in *SIGCOMM*. ACM, 2010, pp. 339–350.

[60] M. Müller-Hannemann and A. Schwartz, "Implementing weighted b-matching algorithms: Insights from a computational study," *ACM Journal of Experimental Algorithmics*, vol. 5, p. 8, 2000.

[61] A. Singla, A. Singh, K. Ramachandran, L. Xu, and Y. Zhang, "Proteus: a topology malleable data center network," in *HotNets*. ACM, 2010.

[62] K. Chen, A. Singla, A. Singh, K. Ramachandran, L. Xu, Y. Zhang, X. Wen, and Y. Chen, "OSA: an optical switching architecture for data center networks with unprecedented flexibility," *IEEE/ACM Trans. Netw.*, vol. 22, no. 2, pp. 498–511, 2014.

[63] K.-T. Foerster, M. Ghobadi, and S. Schmid, "Characterizing the algorithmic complexity of reconfigurable data center architectures," in *ANCS*. IEEE/ACM, 2018.

[64] K.-T. Foerster, M. Pacut, and S. Schmid, "On the complexity of non-segregated routing in reconfigurable data center architectures," *SIGCOMM*, 2019.

[65] T. Fenz, K.-T. Foerster, S. Schmid, and A. Villedieu, "Efficient non-segregated routing for reconfigurable demand-aware networks," in *18th IFIP Networking Conference (IFIP Networking)*, May 2019. [Online]. Available: http://eprints.cs.univie.ac.at/6007/

[66] L. G. Valiant, "A scheme for fast parallel communication," *SIAM J. Comput.*, vol. 11, no. 2, pp. 350–361, 1982.

[67] C. Chang, D. Lee, and Y. Jou, "Load balanced birkhoff-von neumann switches, part I: one-stage buffering," *Comput. Commun.*, vol. 25, no. 6, pp. 611–622, 2002.

[68] W. M. Mellette, R. Das, Y. Guo, R. McGuinness, A. C. Snoeren, and G. Porter, "Expanding across time to deliver bandwidth efficiency and low latency," in *NSDI*, 2020, pp. 1–18.

[69] A. Singla, C. Hong, L. Popa, and P. B. Godfrey, "Jellyfish: Networking data centers randomly," in *NSDI*, 2012. [Online]. Available: https://www.usenix.org/conference/nsdi12/technical-sessions/presentation/singla

[70] Y. Xia, X. S. Sun, S. Dzinamarira, D. Wu, X. S. Huang, and T. S. Eugene Ng, "A tale of two topologies: Exploring convertible data center network architectures with flat-tree," in *SIGCOMM*. ACM, 2017.

[71] G. Porter, R. D. Strong, N. Farrington, A. Forencich, P. Sun, T. Rosing, Y. Fainman, G. Papen, and A. Vahdat, "Integrating microsecond circuit switching into the data center," in *SIGCOMM*. ACM, 2013, pp. 447–458.

[72] A. Goel, M. Kapralov, and S. Khanna, "Perfect matchings in o(nlog n) time in regular bipartite graphs," *SIAM J. Comput.*, vol. 42, no. 3, pp. 1392–1404, 2013.

[73] H. Liu, M. K. Mukerjee, C. Li, N. Feltman, G. Papen, S. Savage, S. Seshan, G. M. Voelker, D. G. Andersen, M. Kaminsky, G. Porter, and A. C. Snoeren, "Scheduling techniques for hybrid circuit/packet networks," in *CoNEXT*. ACM, 2015, pp. 41:1–41:13.

[74] S. B. Venkatakrishnan, M. Alizadeh, and P. Viswanath, "Costly circuits, submodular schedules and approximate carathéodory theorems," *Queueing Syst.*, vol. 88, no. 3-4, pp. 311–347, 2018.

[75] R. Schwartz, M. Singh, and S. Yazdanbod, "Online and offline greedy algorithms for routing with switching costs," *arXiv preprint arXiv:1905.02800*, 2019.

[76] C. Avin, K. Mondal, and S. Schmid, "Demand-aware network designs of bounded degree," in *DISC*, 2017.

[77] ——, "Demand-aware network design with minimal congestion and route lengths," in *Proc. IEEE INFOCOM*, 2019.

[78] C. Avin and S. Schmid, "Toward demand-aware networking: A theory for self-adjusting networks," in *SIGCOMM*, 2018.

[79] C. Avin, A. Hercules, A. Loukas, and S. Schmid, "rdan: Toward robust demand-aware network designs," in *Information Processing Letters (IPL)*, 2018.

[80] S. Anand, N. Garg, and A. Kumar, "Resource augmentation for weighted flow-time explained by dual fitting," in *the twenty-third annual ACM-SIAM symposium on Discrete Algorithms*. SIAM, 2012, pp. 1228–1241.

[81] M. Dinitz and B. Moseley, "Scheduling for weighted flow and completion times in reconfigurable networks," *arXiv preprint arXiv:2001.07784*, 2020.

[82] S. Schmid, C. Avin, C. Scheideler, M. Borokhovich, B. Haeupler, and Z. Lotker, "Splaynet: Towards locally self-adjusting networks," *IEEE/ACM Trans. Netw.*, vol. 24, no. 3, pp. 1421–1433, 2016.

[83] B. Peres, O. Goussevskaia, S. Schmid, and C. Avin, "Concurrent self-adjusting distributed tree networks," in *Proc. International Symposium on Distributed Computing (DISC)*, 2017.

[84] C. Avin, B. Haeupler, Z. Lotker, C. Scheideler, and S. Schmid, "Locally self-adjusting tree networks," in *Proc. 27th IEEE International Parallel and Distributed Processing Symposium (IPDPS)*, May 2013.

[85] B. Peres, O. A. de Oliveira Souza, O. Goussevskaia, C. Avin, and S. Schmid, "Distributed self-adjusting tree networks," in *INFOCOM*. IEEE, 2019.

[86] S. Salman, C. Streiffer, H. Chen, T. Benson, and A. Kadav, "Deepconf: Automating data center network topologies management with machine learning," in *the 2018 Workshop on Network Meets AI & ML*, ser. NetAI'18. New York, NY, USA: ACM, 2018, pp. 8–14. [Online]. Available: http://doi.acm.org/10.1145/3229543.3229554

[87] P. Kalmbach, J. Zerwas, P. Babarczi, A. Blenk, W. Kellerer, and S. Schmid, "Empowering self-driving networks," in *Proc. ACM SIGCOMM 2018 Workshop on Self-Driving Networks (SDN)*, 2018.

[88] T. Truong-Huu, P. M. Mohan, and M. Gurusamy, "Virtual network embedding in ring optical data centers using markov chain probability model," *IEEE Transactions on Network and Service Management*, vol. 16, no. 4, pp. 1724–1738, 2019.

[89] L. Chen, K. Chen, Z. Zhu, M. Yu, G. Porter, C. Qiao, and S. Zhong, "Enabling wide-spread communications on optical fabric with megaswitch," in *NSDI*. Boston, MA: USENIX Association, 2017, pp. 577–593. [Online]. Available: https://www.usenix.org/conference/nsdi17/technical-sessions/presentation/chen

[90] J. Misra and D. Gries, "A constructive proof of vizing's theorem," *Inf. Process. Lett.*, vol. 41, no. 3, pp. 131–133, 1992.

[91] K. Chen, X. Wen, X. Ma, Y. Chen, Y. Xia, C. Hu, Q. Dong, and Y. Liu, "Toward a scalable, fault-tolerant, high-performance optical data center architecture," *IEEE/ACM Transactions on Networking*, vol. 25, no. 4, pp. 2281–2294, 2017.

[92] H. Wang, Y. Xia, K. Bergman, T. S. E. Ng, S. Sahu, and K. Sripanidkulchai, "Rethinking the physical layer of data center networks of the next decade: using optics to enable efficient *-cast connectivity," *Computer Communication Review*, vol. 43, no. 3, pp. 52–58, 2013.

[93] X. S. Huang, X. S. Sun, and T. S. E. Ng, "Sunflow: Efficient optical circuit scheduling for coflows," in *CoNEXT*. ACM, 2016, pp. 297–311.

[94] H. Wang, X. Yu, H. Xu, J. Fan, C. Qiao, and L. Huang, "Integrating coflow and circuit scheduling for optical networks," *IEEE Transactions on Parallel and Distributed Systems*, 2019.

[95] J. Bao, D. Dong, B. Zhao, Z. Luo, C. Wu, and Z. Gong, "Flycast: Free-space optics accelerating multicast communications in physical layer," *Computer Communication Review*, vol. 45, no. 5, pp. 97–98, 2015.

[96] X. S. Sun and T. S. E. Ng, "When creek meets river: Exploiting high-bandwidth circuit switch in scheduling multicast data," in *ICNP*. IEEE Computer Society, 2017, pp. 1–6.

[97] X. S. Sun, Y. Xia, S. Dzinamarira, X. S. Huang, D. Wu, and T. S. E. Ng, "Republic: Data multicast meets hybrid rack-level interconnections in data center," in *ICNP*. IEEE Computer Society, 2018, pp. 77–87.

[98] Y. Xia, T. S. E. Ng, and X. S. Sun, "Blast: Accelerating high-performance data analytics applications by optical multicast," in *INFOCOM*. IEEE, 2015, pp. 1930–1938.

[99] L. Luo, K. Foerster, S. Schmid, and H. Yu, "Splitcast: Optimizing multicast flows in reconfigurable datacenter networks," in *INFOCOM*. IEEE, 2020.

[100] S. B. Venkatakrishnan, M. Alizadeh, and P. Viswanath, "Costly circuits, submodular schedules and approximate carathéodory theorems," in *SIGMETRICS*. ACM, 2016, pp. 75–88.

[101] S. Gringeri, N. Bitar, and T. J. Xia, "Extending software defined network principles to include optical transport," *IEEE Communications Magazine*, vol. 51, no. 3, pp. 32–40, March 2013.

[102] K. Christodoulopoulos, P. Kokkinos, and E. M. Varvarigos, "Indirect and direct multicost algorithms for online impairment-aware rwa," *IEEE/ACM Transactions on Networking*, vol. 19, no. 6, pp. 1759–1772, 2011.

[103] J. Oliveira, A. Caballero, E. M. aes, U. Moura, R. Borkowski, G. Curiel, A. Hirata, L. Hecker, E. Porto, D. Zibar, J. M. ao, I. T. Monroy, and J. Oliveira, "Demonstration of edfa cognitive gain control via gmpls for mixed modulation formats in heterogeneous optical networks," in *Optical Fiber Communication Conference/National Fiber Optic Engineers Conference 2013*. Optical Society of America, 2013, p. OW1H.2. [Online]. Available: http://www.osapublishing.org/abstract.cfm?URI=OFC-2013-OW1H.2

[104] U. Moura, M. Garrich, H. Carvalho, M. Svolenski, A. Andrade, A. C. Cesar, J. Oliveira, and E. Conforti, "Cognitive methodology for optical amplifier gain adjustment in dynamic dwdm networks," *Journal of Lightwave Technology*, vol. 34, no. 8, pp. 1971–1979, 2016.

[105] U. Moura, M. Garrich, A. C. Cesar, J. Oliveira, and E. Conforti, "Execution time improvement for optical amplifier cognitive methodology in dynamic wdm networks," in *2017 SBMO/IEEE MTT-S International Microwave and Optoelectronics Conference (IMOC)*. IEEE, 2017, pp. 1–5.

[106] J. Case, M. Fedor, M. L. Schoffstall, and J. Davin, "Rfc1157: Simple network management protocol (snmp)," 1990.

[107] R. Enns, M. Bjorklund, J. Schoenwaelder, and A. Bierman, "Network configuration protocol (netconf)," 2011.

[108] A. Shaikh, T. Hofmeister, V. Dangui, and V. Vusirikala, "Vendor-neutral network representations for transport sdn," in *2016 Optical Fiber Communications Conference and Exhibition (OFC)*, 2016, pp. 1–3.

[109] A. V. Lehmen, R. Doverspike, G. Clapp, D. M. Freimuth, J. Gannett, K. Kim, H. Kobrinski, E. Mavrogiorgis, J. Pastor, M. Rauch, K. K. Ramakrishnan, R. Skoog, B. Wilson, and S. L. Woodward, "Coronet: Testbeds, cloud computing, and lessons learned," in *Optical Fiber Communication Conference*. Optical Society of America, 2014, p. W4B.1. [Online]. Available: http://www.osapublishing.org/abstract.cfm?URI=OFC-2014-W4B.1

[110] R. A. Skoog and A. L. Neidhardt, "A fast, robust signaling protocol for enabling highly dynamic optical networks," in *2009 Conference on Optical Fiber Communication-incudes post deadline papers*. IEEE, 2009, pp. 1–3.

[111] H. Zang, J. P. Jue, B. Mukherjee *et al.*, "A review of routing and wavelength assignment approaches for wavelength-routed optical wdm networks," *Optical networks magazine*, vol. 1, no. 1, pp. 47–60, 2000.

[112] A. S. Thyagaturu, A. Mercian, M. P. McGarry, M. Reisslein, and W. Kellerer, "Software defined optical networks (sdons): A comprehensive survey," *IEEE Commun. Surv. Tutorials*, vol. 18, no. 4, pp. 2738–2786, 2016.

[113] R. Casellas, R. Martínez, R. Vilalta, and R. M. noz, "Control, management, and orchestration of optical networks: Evolution, trends, and challenges," *J. Lightwave Technol.*, vol. 36, no. 7, pp. 1390–1402, Apr 2018. [Online]. Available: http://jlt.osa.org/abstract.cfm?URI=jlt-36-7-1390

[114] M. Kodialam and T. V. Lakshman, "Integrated dynamic ip and wavelength routing in ip over wdm networks," in *IEEE INFOCOM*, vol. 1, 2001, pp. 358–366.

[115] A. Brzezinski and E. Modiano, "Dynamic reconfiguration and routing algorithms for ip-over-wdm networks with stochastic traffic," *Journal of Lightwave Technology*, vol. 23, no. 10, p. 3188, 2005.

[116] Z. Zhong, N. Hua, M. Tornatore, J. Li, Y. Li, X. Zheng, and B. Mukherjee, "Provisioning short-term traffic fluctuations in elastic optical networks," *IEEE/ACM Trans. Netw.*, vol. 27, no. 4, pp. 1460–1473, 2019.

[117] S. K. Patri, A. Autenrieth, D. Rafique, J.-P. Elbers, and C. M. Machuca, "Hecson: Heuristic for configuration selection in optical network planning," in *Optical Fiber Communication Conference (OFC) 2020*. Optical Society of America, 2020, p. Th2A.32. [Online]. Available: http://www.osapublishing.org/abstract.cfm?URI=OFC-2020-Th2A.32

[118] D. Xu, G. Li, B. Ramamurthy, A. L. Chiu, D. Wang, and R. D. Doverspike, "On provisioning diverse circuits in heterogeneous multi-layer optical networks," *Comput. Commun.*, vol. 36, no. 6, pp. 689–697, 2013.

[119] C. Mas, I. Tomkos, and O. K. Tonguz, "Failure location algorithm for transparent optical networks," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 8, pp. 1508–1519, 2005.

[120] J. Chen, L. Wosinska, C. M. Machuca, and M. Jaeger, "Cost vs. reliability performance study of fiber access network architectures," *IEEE Communications Magazine*, vol. 48, no. 2, pp. 56–65, 2010.

[121] L. Luo, H. Yu, K.-T. Foerster, M. Noormohammadpour, and S. Schmid, "Inter-datacenter bulk transfers: Trends and challenges," *IEEE Network*, vol. to appear, 2020.

[122] X. Jin, Y. Li, D. Wei, S. Li, J. Gao, L. Xu, G. Li, W. Xu, and J. Rexford, "Optimizing bulk transfers with software-defined optical WAN," in *SIGCOMM*, 2016. [Online]. Available: http://doi.acm.org/10.1145/2934872.2934904

[123] X. Jin, H. H. Liu, R. Gandhi, S. Kandula, R. Mahajan, M. Zhang, J. Rexford, and R. Wattenhofer, "Dynamic scheduling of network updates," in *SIGCOMM*. ACM, 2014, pp. 539–550.

[124] L. Luo, K. Foerster, S. Schmid, and H. Yu, "Deadline-aware multicast transfers in software-defined optical wide-area networks," *IEEE Journal on Selected Areas in Communications*, 2020.

[125] S. Jia, X. Jin, G. Ghasemiesfeh, J. Ding, and J. Gao, "Competitive analysis for online scheduling in software-defined optical wan," in *INFOCOM*, 2017.

[126] M. Dinitz and B. Moseley, "Scheduling for weighted flow and completion times in reconfigurable networks," in *INFOCOM*, 2020.

[127] R. Singh, M. Ghobadi, K. Foerster, M. Filer, and P. Gill, "Run, walk, crawl: Towards dynamic link capacities," in *HotNets*. ACM, 2017.

[128] R. Singh, M. Ghobadi, K.-T. Foerster, M. Filer, and P. Gill, "Radwan: rate adaptive wide area network," in *the 2018 Conference of the ACM Special Interest Group on Data Communication*. ACM, 2018, pp. 547–560.

[129] K. Foerster, S. Schmid, and S. Vissicchio, "Survey of consistent software-defined network updates," *IEEE Communications Surveys and Tutorials*, vol. 21, no. 2, pp. 1435–1461, 2019.

[130] S. Tseng, "Perseverance-aware traffic engineering in rate-adaptive networks with reconfiguration delay," in *ICNP*. IEEE, 2019, pp. 1–10.

[131] K.-T. Foerster, L. Luo, and M. Ghobadi, "Optflow: A flow-based abstraction for programmable topologies," in *SOSR*. ACM, 2020.

[132] Y. Wang, Z. McNulty, and H. Nguyen, "Network virtualization in spectrum sliced elastic optical path networks," *Journal of Lightwave Technology*, vol. 35, no. 10, pp. 1962–1970, 2017.

[133] P. Soto, P. Maya, and J. F. Botero, "Resource allocation over eon-based infrastructures in a network virtualization environment," *IEEE Transactions on Network and Service Management*, vol. 16, no. 1, pp. 13–26, 2019.

[134] A. E. Ozdaglar and D. P. Bertsekas, "Routing and wavelength assignment in optical networks," *IEEE/ACM Transactions on Networking*, vol. 11, no. 2, pp. 259–272, 2003.

[135] P. Kokkinos, P. Soumplis, and E. A. Varvarigos, "Pattern-driven resource allocation in optical networks," *IEEE Transactions on Network and Service Management*, vol. 16, no. 2, pp. 489–504, 2019.

[136] R. Singh, M. Ghobadi, K.-T. Foerster, M. Filer, and P. Gill, "Radwan: Rate adaptive wide area network," in *the 2018 Conference of the ACM Special Interest Group on Data Communication*, ser. SIGCOMM '18. New York, NY, USA: ACM, 2018, pp. 547–560. [Online]. Available: http://doi.acm.org/10.1145/3230543.3230570

[137] N. Benzaoui, M. S. Gonzalez, J. M. Estarán, H. Mardoyan, W. Lautenschlaeger, U. Gebhard, L. Dembeck, S. Bigo, and Y. Pointurier, "Deterministic dynamic networks (ddn)," *Journal of Lightwave Technology*, vol. 37, no. 14, pp. 3465–3474, 2019.

[138] C.-Y. Hong, S. Kandula, R. Mahajan, M. Zhang, V. Gill, M. Nanduri, and R. Wattenhofer, "Achieving high utilization with software-driven wan," *SIGCOMM*, vol. 43, no. 4, pp. 15–26, 2013.

[139] S. Figueira, S. Naiksatam, H. Cohen, D. Cutrell, P. Daspit, D. Gutierrez, D. B. Hoang, T. Lavian, J. Mambretti, S. Merrill *et al.*, "Dwdm-ram: Enabling grid services with dynamic optical networks," in *IEEE International Symposium on Cluster Computing and the Grid, 2004. CCGrid 2004*. IEEE, 2004, pp. 707–714.

[140] E. Bernier, M. Vukovic, D. Goodwill, P. Daspit, and G. Wang, "Omninet: A metropolitan 10 gb/s dwdm photonic switched network trial," in *Optical Fiber Communication Conference*. Optical Society of America, 2004, p. WH4.

[141] P. Iovanna, R. Sabella, and M. Settembre, "A traffic engineering system for multilayer networks based on the gmpls paradigm," *IEEE Network*, vol. 17, no. 2, pp. 28–37, 2003.

[142] M. Chamania, M. Caria, and A. Jukan, "Achieving ip routing stability with optical bypass," *Optical Switching and Networking*, vol. 7, no. 4, pp. 173–184, 2010. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1573427710000305

[143] X. Jin, Y. Li, D. Wei, S. Li, J. Gao, L. Xu, G. Li, W. Xu, and J. Rexford, "Optimizing bulk transfers with software-defined optical wan," in *the 2016 ACM SIGCOMM Conference*. ACM, 2016, pp. 87–100.

[144] J. Gossels, G. Choudhury, and J. Rexford, "Robust network design for ip/optical backbones," *IEEE/OSA Journal of Optical Communications and Networking*, vol. 11, no. 8, pp. 478–490, 2019.

[145] S. K. Mani, M. Nance Hall, R. Durairajan, and P. Barford, "Characteristics of metro fiber deployments in the us," in *the Network Traffic Measurement and Analysis Conference*, 2020.

[146] S. Oda, M. Miyabe, S. Yoshida, T. Katagiri, Y. Aoki, J. C. Rasmussen, M. Birk, and K. Tse, "A learning living network for open roadm networks," in *ECOC 2016; 42nd European Conference on Optical Communication*. VDE, 2016, pp. 1–3.

[147] P. Berde, M. Gerola, J. Hart, Y. Higuchi, M. Kobayashi, T. Koide, B. Lantz, B. O'Connor, P. Radoslavov, W. Snow *et al.*, "Onos: towards an open, distributed sdn os," in *the third workshop on Hot topics in software defined networking*, 2014, pp. 1–6.

[148] M. Nance Hall, G. Liu, R. Durairajan, and V. Sekar, "Fighting fire with light: Tackling extreme terabit ddos using programmable optics," in *the Workshop on Secure Programmable Network Infrastructure*, ser. SPIN '20. New York, NY, USA: Association for Computing Machinery, 2020, p. 42–48. [Online]. Available: https://doi.org/10.1145/3405669.3405824

[149] J. M. Smith and M. Schuchard, "Routing around congestion: Defeating ddos attacks and adverse network conditions via reactive bgp routing," in *2018 IEEE Symposium on Security and Privacy (SP)*. IEEE, 2018, pp. 599–617.

[150] S. K. Fayaz, Y. Tobioka, V. Sekar, and M. Bailey, "Flexible and elastic ddos defense using bohatei," in *USENIX Security*, 2015.

[151] M. Zhang, G.-Y. Li, S. Wang, C. Liu, A. Chen, H. Hu, G. Gu, Q. Li, M. Xu, and J. Wu, "Poseidon: Mitigating volumetric ddos attacks with programmable switches," 2020.

[152] C. Avin, M. Ghobadi, C. Griner, and S. Schmid, "On the complexity of traffic traces and implications," *Proc. ACM Meas. Anal. Comput. Syst.*, vol. 4, no. 1, pp. 20:1–20:29, 2020.