On Efficient Oblivious Wavelength Assignments for Programmable Wide-Area Topologies

Thomas Fenz
University of Vienna
Faculty of Computer Science
Austria
thomas.fenz@univie.ac.at

Klaus-Tycho Foerster
TU Dortmund
Germany
klaus-tycho.foerster@tu-dortmund.de

Stefan Schmid
TU Berlin, University of Vienna, Fraunhofer SIT
Germany and Austria
stefan.schmid@tu-berlin.de

ABSTRACT

Given the explosively growing traffic related to data-centric applications and AI, especially to and from the cloud, it is crucial to make the best use of the given resources of wide-area backbone networks (WANs). An intriguing approach to improve both efficiency and performance of WANs is to render networks more adaptive and "demand-aware", on the physical layer: innovative programmable wide-area topologies support dynamic wavelength assignments. This is enabled by the application of colorless and directionless Reconfigurable Optical Add/Drop Multiplexers (CD ROADM), and by leveraging the capabilities of software-defined controllers.

This paper investigates the benefit of such fully dynamic wavelength assignments in programmable WAN topologies, compared to an oblivious wavelength assignment. To this end, we also propose a new demand-oblivious strategy to optimize the capacity of a WAN. Considering both real and synthetic scenarios, we find that our proposed demand-oblivious strategy can perform close to dynamic approaches with respect to throughput, without entailing reconfiguration costs.

CCS CONCEPTS

• Networks → Network architectures; **• Theory of computation → Design and analysis of algorithms**.

KEYWORDS

wide area networks, optical reconfiguration, network design

ACM Reference Format:
Thomas Fenz, Klaus-Tycho Foerster, and Stefan Schmid. 2021. On Efficient Oblivious Wavelength Assignments for Programmable Wide-Area Topologies. In *Symposium on Architectures for Networking and Communications Systems (ANCS '21), December 13–16, 2021, Layfette, IN, USA.* ACM, New York, NY, USA, 13 pages. https://doi.org/10.1145/3493425.3502753

1 INTRODUCTION

With the increasing popularity of distributed and data-centric applications (e.g., related to business, social networking, or health) and artificial intelligence, network traffic is currently growing explosively, especially to and from datacenters (i.e., the cloud) [10, 32, 48, 50, 69, 73, 79]. Novel applications, such as emerging wide-area analytics platforms running data-parallel jobs across geo-distributed sites [62], are likely to increase traffic further in the near future. To meet the increasingly stringent requirements on the efficiency and performance of wide-area backbone networks (WAN), Internet Service Providers (ISP) hence spend billions of dollars on enhancing their infrastructure by building access points across the world and interconnect them using high-speed fiber optics.

An intriguing approach to render networks more efficient and performant, is to make them reconfigurable and "adaptive". Especially the notion of software-defined WANs, using modern remote software controllers, have recently received much attention to improve networks [30, 31, 34, 35, 38]. Wavelength division multiplexing (WDM) devices such as colorless and directionless Reconfigurable Optical Add-Drop Multiplexers (CD ROADMs) [3, 14, 54, 65, 67] further enable on-demand reconfiguration of the physical topology. With the wavelength selective switching (WSS) components in ROADMs, any signal on an input port can be switched to any output port on a per-wavelength basis. Next-generation ROADMs are capable of performing such reconfigurations in a fraction of a second [9, 33, 36, 43]. Due to advances achieved with ROADMs, operators are hence no longer constrained to plan the entire wavelength assignment long term.

The potential benefits of reconfigurability have already been demonstrated in the context of data centers networks (DCNs) [8, 15, 16, 20, 22, 25, 39, 43–45, 49, 57, 59, 83]. In datacenter (DC) networks, nodes can be connected directly via wireless/optical links or via optical switches to increase

ANCS '21, December 13–16, 2021, Layfette, IN, USA

ACM ISBN 978-1-4503-9168-9/21/12.
https://doi.org/10.1145/3493425.3502753

the overall network performance. Although inspiring, by considering the variety and the geographical scale of existing WAN topologies and its implications regarding reconfigurability, a simple adaption of proposed solutions seems infeasible.

This paper hence studies the potential benefits of reconfigurable WANs, when jointly optimizing the flexibilities introduced by both topology programming (TP) and traffic engineering (TE), as proposed in the literature [17, 29, 36, 46, 53].

Unlike prior work, which investigated the efficient joint optimization of TP+TE in the WAN, we tackle the problem from the opposite direction: We investigate how to design optimized *demand-oblivious* wavelength allocations and study their performance in WAN topologies. Contrary to joint optimization of TP+TE, in the demand-oblivious approach, the capacity-design does not depend on any assumptions about the traffic carried by the network. Demand-oblivious designs only need to be applied once in advance and are as thus attractive as they avoid reconfiguration overheads and can hence be more scalable, as previously demonstrated on the routing layer [41, 64].

Contributions: Motivated by the fast growth of network traffic to and from the cloud, in this work, we take the next steps towards designing efficient demand-oblivious network capacity allocations for the WAN. Our main contributions are twofold:

(1) We show that already a simple oblivious wavelength allocation achieves on average within 8% of the throughput of an optimal TP+TE formulation in the WAN, avoiding the downsides of physical reconfiguration. Our results are based on an evaluation of 169 WANs from the Topology Zoo data set, utilizing synthetic traffic matrices. To validate our approach, we also compare the throughput benefits of the synthetic traffic with real-world data for two WANs, where the median benefits closely match for both synthetic and real traffic.
(2) Second, we show that for a small set of WAN topologies, the benefits of joint TP+TE can be much larger, up to 40%. Focusing on these highly optimizable topologies, we present an improved demand-oblivious network capacity design strategy. Inspired by datacenter topology design, we reduce the all-to-all communication bottlenecks, which in turn significantly reduces the gap towards joint demand-aware TP+TE formulations to below 5%, on average. On the largest third of the evaluated topologies, this gap is halved.

Overview: In Section 2, we introduce the core components in reconfigurable WANs and topology programming. We continue with a description of our model and the underlying assumptions in Section 3, followed by an Integer Linear Programming (ILP) TP+TE problem formulation in Section 4. Next in Section 5, we evaluate the performance of TP+TE against oblivious, namely uniform, capacity designs on a large data set of real-world WAN topologies, concluding that in general, the throughput benefits are relatively small. However, some topologies allow larger gains, and we show in Section 6 how more refined oblivious strategies can significantly reduce this gap. We discuss related work in Section 7 and lastly conclude in Section 8.

2 BACKGROUND

In this section, we first introduce the core components in modern optical WANs and briefly discuss the shift from static network design to dynamic topology programming. We refer to a recent survey [28] for a general overview.

IP/Optical WAN backbones. Modern WAN backbones comprise edge nodes where traffic enters and exits the network, optical fibers, IP nodes (ROADMs), connecting the fiber and optical nodes to regenerate the optical signal over long distances [23, 24]. Traffic is caused by end systems outside the backbone and aggregated on each edge node. IP nodes consist of tails and regenerators, where a tail combines a transponder to convert optical signals to electrical signals and a router port connected to a router. Regenerators (regens) in IP nodes as well as in optical nodes pass through (amplified) optical signals without conversion.

This results in a two-layer view of the network: the IP layer where traffic engineering is applied and the underlying optical layer. By leveraging the capabilities of software-defined controllers, we can jointly optimize the IP layer and the underlying optical layer. By reprogramming both, the CD ROADMs and the routers, we can change the physical topology on the fly and so further optimize routing. E.g. transponders and router ports at a single IP Node can be combined to build a faster link. The potential of centralized control is seen in Microsoft's SWAN [31], where the network utilization is significantly improved. Another example of a successful application is Google's B4 [30, 34]. On a five-year scale (2013-2018) the traffic volume in Google's private WAN increased by a hundredfold, at the same time the availability was improved from 99% to 99.99% and reveal the potential of a global software-defined WAN at a massive scale.

Reconfigurable Optical Add Drop Multiplexer (ROADMs). Colorless directionless Reconfigurable Optical Add Drop Multiplexer (CD ROADMs) are the crucial devices in optical WAN topologies. CD ROADMs enable an all-optical transmission of light signals from source to destination and avoid the need for expensive optical-electronic-optical (O-E-O) signal conversions. In Fig. 1 a 4-Degree ROADM is shown. Utilizing the wavelength selective switching (WSS) component, ROADMs allow steering light signals from any input port to any output port on a per-wavelength basis. Furthermore, they allow to add or drop wavelengths to or from

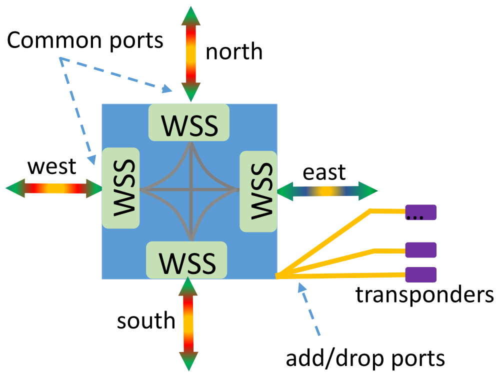
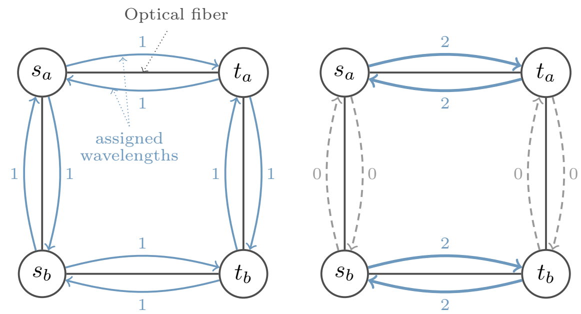


Figure 1: 4-Degree ROADM [21]



(a) Uniform wavelength assignment (baseline)

(b) Reconfigured wavelength assignment (TP)

Figure 2: Throughput optimization by means of topology programming (TP).

network nodes without interfering with other wavelengths [54, 58, 65]. ROADMs allow dynamic software-driven monitoring and enable to seamlessly change connectivity as needed on the fly. First-generation ROADMs were limited to fixed wavelength assignments mapped to each add/drop port, and a reconfiguration needed to be done by technicians on site. Colorless ROADMs extend this technology by enabling the automation of the assignment of add/drop wavelength functionality. Directionless ROADMs allow any wavelength to be switched to any direction, i.e port in a software-defined manner. While in today's ROADMs these functionalities are widely employed, adding support for contentionless switching technologies (CDC ROADMs) maximizes the flexibility at the optical layer. Next-generation ROADMs support to transmit and receive the same wavelength in and from multiple directions at the same time [3, 14].

Topology Programming (TP). The broad employment of ROADMs in modern WAN architectures leads to a paradigmatic change in network design. While the placement of hardware components must still be carefully planned in advance, now, adjusting the configuration of such optical switches enables to reconfigure the network topology. By changing the wavelength assignment of ROADMs, in other words, assigning wavelengths to connected fiber, we can change the optical topology and as such the capacity of edges in the order of seconds to minutes, depending on the circumstances and the further needed hardware adaptions, e.g., for amplifiers on long-range links. A change in the wavelength assignment of ROADMs enables us to dynamically establish new IP links to what we further refer to as topology programming (TP).

Note that in WANs the duration of such reconfigurations are orders of magnitudes slower than in datacenters, and that during reconfiguration, no data can be sent over the wavelengths [71]. While there is further promising research on innovative hardware [27, 82], currently, programming the wavelengths beyond metro use cases incurs relatively long delays, and hence comes at an inherent trade-off. We refer to this as the *reconfiguration cost*.

3 MODEL

Before we evaluate the potential of reconfigurability, we discuss our model abstraction and the underlying assumptions.

Nodes in our model represent ROADMs, acting at the same time as the Points of Presence (PoP) of the WAN topology. Of course, traffic originates and terminates outside the optical WAN topology, but our optimizations are performed based on a snapshot of the aggregated demand at the edge of the network, i.e, each node in our model. The links in the network topology are optical fiber. Gossel [23] points out that not the fiber between these PoPs limit the network's overall performance, but ROADMs restricted in the number of optical ports are the limiting factor.

Single Logical Controller. We assume the presence of a single logical remote controller, which receives any change in traffic demands between any pair of nodes. With this information, combined with the information of the underlying optical network, we can compute the optimal solution and further, utilize the controller to remotely reconfigure the wavelength assignment of the ROADMs (TP) and provide optimal routing paths for the current traffic (TE).

Bidirectional wavelengths. Recent work [36, 46] shows the potential of programmable topologies in respect to optimizing efficiency, but assume a unidirectional wavelength assignment. We extend their model addressing current developments in optical switching technologies. Vendors of ROADMs often bundle sending and receiving components in pairs, which we take into account by adding the restriction of bidirectionality. So, a wavelength assignment $(v_i, v_j) \rightarrow w_{ij}$ implies an assignment $(v_j, v_i) \rightarrow w_{ji}$ with $w_{ij} = w_{ji}$.

Example. In Fig. 2 we give an intuition of our IP/optical network abstraction. Each of the four ROADMs (nodes) supports two pairs of wavelengths, i.e. two incoming and two outgoing information streams. In this example, a flow rate of 2 from both $s_a \rightarrow t_a$ and from $s_b \rightarrow t_b$ is requested. In Fig. 2a each ROADM assigns the wavelengths equally to its adjacent links, which we further refer to as uniform capacity design. With this configuration, a maximum flow rate of 1 can be processed for both requested flows. The optimal reassignment of wavelengths is shown in Fig. 2b, where the wavelength shift results in doubling the overall throughput.

4 ILP FORMULATION

The fundamental idea is to optimize the objective by reconfiguring the optical layer, i.e., assigning wavelengths to links to maximize the overall throughput. To this end, we present an ILP to optimize the joint TP+TE problem and follow by decoupling the TE from the joint optimization formulation.

Joint TP+TE ILP. Given a network topology, D_{st} is the demanded flow size from node s to node t. The variable w_{ij} represents the number of wavelengths assigned to the link from i to j, and the variable f_{st} is the fraction of the flow from s to t the network is capable to route. The number of wavelengths a node i can use to send/receive is defined by $deg_i \cdot \frac{\beta}{\alpha}$, where deg_i signifies the degree of the node i and β is the number of wavelengths an edge can carry. α is a constant greater than 1, used to model that interfaces of a node (ROADM) have limited ports, and a node can not use all wavelengths of the connected edges. γ denotes the capacity of a single wavelength. The variable x_{ij}^{st} denotes the fraction of the flow from s to t that traverses the link from i to j.

Objective: The goal is to maximize the overall throughput.

$$\max \sum_s \sum_t D_{st} f_{st} \tag{1}$$

Flow conservation: A splittable flow that enters a node must leave it, except for the source and destination.

$$\sum_j x_{ij}^{st} - \sum_j x_{ji}^{st} = \begin{cases} f_{st}, & \text{if } i = s. \\ -f_{st}, & \text{if } i = t. \\ 0, & \text{otherwise.} \end{cases} \tag{2}$$

Capacity constraint: The sum of flows on a link must not exceed its capacity ($w_{ij}\gamma$).

$$\sum_s \sum_t x_{ij}^{st} D_{st} \leq w_{ij}\gamma \tag{3}$$

Nodes limit: A node can process only a limited number of wavelengths.

$$\sum_j w_{ij} = deg_i \cdot \frac{\beta}{\alpha} \tag{4}$$

Bidirectionality: Each wavelength must be assigned in both directions.

$$w_{ij} = w_{ji} \tag{5}$$

Max wavelengths on link: No link can carry more than β wavelengths.

$$w_{ij} \leq \beta \tag{6}$$

TE LP. To decouple the TE from the joint optimization formulation, we only use (1)-(3). Additionally, w, the wavelength assignment, is passed as an input parameter. As w is the only integer variable in the Joint TP+TE ILP, the program solving the traffic engineering problem is hence an LP.

5 EVALUATING JOINT TP+TE

Dynamically adjusting the optical layer of the WAN topology according the current demand can improve the maximum overall throughput of the network. To evaluate the full potential of such reconfigurations on real-world WAN topologies, we conducted extensive simulations in which we show the gain of a joint TP+TE optimization compared to a TE optimization with a static uniform wavelength assignment.

In this section, we first briefly discuss the test environment and elaborate on the parameters regarding reconfigurability. We then present a demand-oblivious uniform capacity design followed by the relaxation of the Joint TP+TE ILP. We review the real-world data used for our simulations and give a small study of the suitability of synthetic traffic generated with the Gravity Model (GM). To this end, we compare the TP+TE gains under the GM with the benefits under real-world traffic data from SNDLib [55, 56, 72].

We conclude this section with an evaluation of all 169 networks available at TopologyZoo [40, 75] with $|V| \leq 100$.

Our evaluation shows that already the simple uniform design achieves on average within 8% of the throughput of an optimal TP+TE formulation, however that for some WAN topologies, the benefits can be much larger, up to 40%. This brings us to further investigations in the subsequent section.

5.1 Methodology

Test Environment. All simulations were executed on an HP DL380 G9 with 2x Intel Xeons E5-2697V3 SR1XF with 2.6 GHz, 14 cores each, and a total of 128 GB DDR4 RAM.

The host machine was running Ubuntu 18.04.4 LTS. We implemented the proposed algorithms in Python (3.7)[63] leveraging the libraries NetworkX (2.4)[51], Numpy (1.18)[52], and SciPy (1.1)[68]. To solve the LPs we used Gurobi (9.0)[26].

Optics Parameter. Modern fiber is capable of carrying up to 100 wavelengths or even more. In our model, this is denoted by the parameter β which we fixed to 100 in all our experiments. In reality, it is often not the fiber which limits the wavelengths to be utilized between two nodes, but rather the interface of the optical hardware. To express this in our model and to tune our experiments interestingly, we limit the node's capability of processing wavelengths as shown in Equation (4) in Section 4 by $deg_i \cdot \frac{\beta}{\alpha}$. With α, we model the fact that interfaces of ROADMs are limited in the number of ports. When setting α to 1.0, each fiber can be completely filled up to its maximum of β wavelengths and no further capacity optimization is possible. When setting α to larger values, the algorithms have to decide how to distribute the wavelengths over the nodes' neighbors—and the uniform approach simply assigns the same value of $\frac{\beta}{\alpha}$ everywhere. α is also an theoretical upper bound factor of the optimization potential compared to the setting with a uniform wavelength assignment: if each node had α times more wavelengths, each fiber would be completely filled with wavelengths. In our experiments, we fix the value of α to 2, which suffices since the maximum improvement of joint TP+TE over Uniform-TP never comes within reach of this limit. The bandwidth of an optical fiber link depends on the number of wavelengths assigned to it and the capacity of each wavelength carrying information which we denote by γ. Modern optical technology supports transfer rates of over 100 GB/s. We set the parameter γ to 100 in all our experiments.

Uniform Capacity Design. In order to obtain a first and natural demand-oblivious capacity design, we propose that each node spreads its wavelengths equally over all neighbors. An simple example is given in Fig. 2a, where each node can support two wavelengths and has two neighbors, in turn assigning one wavelength to each link. More generally, the wavelength assignment w in uniform capacity design is defined by $w_{ij} = \beta/\alpha \ \ \forall (i,j) \in E$. The resulting wavelength assignment serves as an input of the TE algorithm.

Relaxation of the Joint TP+TE ILP. Due to the complexity of the ILP introduced in Section 4, we model the problem in an LP formulation to support extensive simulations on a large set of topologies. For this, we relax the integrality of the wavelength assignment w, which results into a slightly above-optimal solution. With respect to comparisons, this puts our oblivious wavelength assignments at a small disadvantage, i.e., will we perform better in practice, as the relaxation gives a true upper-bound on achievable throughput. In other words, the relaxation puts the joint TP+TE approach (our baseline comparison/"competitor") at a (slight) advantage.

Data Sources. For our simulations, we used real-world traffic in the form of accumulated traffic matrices and the related topology data taken from SNDLib [55, 56, 72]. The second source of real-world data is TopologyZoo [40, 75] which provide a large body of structural information of topologies. To vary the input data regarding traffic, we synthesized TMs which we further mapped to all used topologies.

Traffic Matrices (TM). TE algorithms, but also of the underlying network design task of the joint TP+TE algorithm, require information about the demanded traffic to be served. The most important representation is the traffic matrix (TM).

A simple and efficient way to synthesize TMs is proposed by Roughan [66]: the Gravity Model (GM). Its name derives from Newton's law of gravitation and it is applied in many fields of science [60, 61, 74]. In the GM it is assumed that each traffic flow from a source to a destination node is independent of other packets. In the context of WAN backbones, they are ingress/egress nodes, i.e., edge switches.

We implemented the Gravity Model as proposed in [66] to synthesize TMs and map the corresponding demands to the nodes of the WAN topology.

5.2 Evaluating the Optimization Potential

As a demand-oblivious baseline, we evaluate the maximum overall throughput on a uniform capacity design with the proposed TE LP introduced in Section 4. We then compare the results gained with the joint TP+TE LP, providing an upper bound for any throughput under reconfiguration, and evaluate the potential of reconfigurability of a large set of topologies. To investigate the critical traffic distributions, we scaled the traffic flows such that 70-90% of the requested flow can be served.[1]

Suitability of Synthetic Traffic. The proprietary nature of ISPs and backbones results in a lack of real-world traffic data. Besides traffic matrices extracted from scientific research networks, we augment our simulations by artificially generated traffic matrices to enable the evaluation of arbitrary topologies. In Fig. 3, we illustrate the requested flow rates in a series of heat maps, comparing the traffic data for the Abilene network available from SNDLib and synthetically generated TMs using the GM. We scaled all demands of a single TM by a factor such that the sums of all TMs are equal. Rows having many red (blue) squares reflect high (low) amounts of traffic entering the network at this node. Analogously, columns comprising many red (blue) squares reflect high (low) amounts of traffic exiting the network at this node. We observe that both, real and synthetic traffic, show spatial locality, i.e., rows/columns highlighted by many red, respectively, many blue squares. Furthermore, whereas the real traffic shows also temporal locality, i.e., the rows and columns show similarities over different times the traffic was captured, each synthetic TM is generated using different randomly generated weight vectors independent of previously generated TMs, so all temporal information is eliminated.

[1] If the network is heavily underutilized, even a naïve oblivious capacity design has identical throughput as joint TP+TE, analogously for heavy over-utilizations. We hence investigate traffic that can (barely) still be served by joint TP+TE, in order to study the full potential of such joint optimizations.

In Fig. 4, we show an evaluation of reconfigurability on the Abilene and the Géant network using real and synthetic traffic. For the experiments on Abilene/Géant with real-world data, we used TMs and the topology data from SNDLib [55, 56, 72] each aggregated over 5 min/15 min and equally spread within a day. The absence of temporal locality using synthetic TMs results in a higher variance in benefits of reconfigurability.

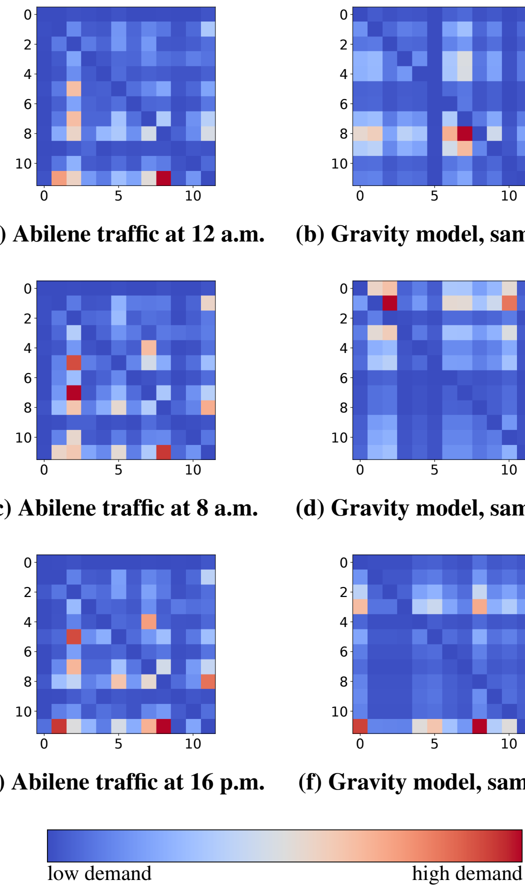
Using this model to generate TMs provides sufficient uncertainty to apply the TMs as input for networking algorithms to evaluate.

Since we ignore reconfiguration costs in this work, we can evaluate traffic based on independent TMs without modeling temporal locality. The relative difference is already small for our sample size of 10 matrices and we expect it to be smaller with still more samples. We thus conclude that the investigation of synthetic GM traffic matrices is a reasonable assumption our further study of real-world topologies.

Evaluation of TopologyZoo Data. In a next step we evaluated topology data available at TopologyZoo [40, 75]. For our experiments, we restrict the network size to $|V| \leq 100$ and only evaluate the newest version of each topology, resulting in 169 topologies. In Fig. 5, we show the joint TP+TE reconfigurability gains of each topology, tested with 10 randomly generated TMs using the GM. All in all we can see that the benefits of joint TP+TE over a uniform wavelength capacity allocation are low, in average exhibiting just 7.85% further throughput potential, over all topologies. On the other hand, we see that some selected topologies allow for larger gains, which we study further next. Different to data centers, where networks are clearly and uniformly structured, e.g. in Clos topologies, WANs commonly develop and expand over time and are hence unique in their topologies. To provide some intuition on our results, in Fig. 6, we visualized three sample networks with a Kamada-Kawai force-directed layout [37].

The high connectivity of the Chinanet topology preserves any optimization through TP, as the uniform wavelength assignment already exhibits good conditions for optimizing the throughput with TE only. We observe long chains of 2-degree-nodes dominating the Lambdanet topology, without much potential for throughput gains by means of topology programming. Contrary to that, the Sinet topology has a high optimization potential, as it has neither the simple structure of Lambdanet nor the high connectivity of Chinanet.

A uniform capacity design misses the separation between peripheral nodes, requiring less capacity, and the sparse backbone links, requiring relatively large capacity. Here TP can leverage a strong core interconnect, leading to a gain in throughput of over 30% in average. We investigate next how such throughput potential can also be realized by oblivious designs.



(a) Abilene traffic at 12 a.m. **(b) Gravity model, sample 1**

(c) Abilene traffic at 8 a.m. **(d) Gravity model, sample 2**

(e) Abilene traffic at 16 p.m. **(f) Gravity model, sample 3**

Figure 3: Comparison of real traffic from Abilene network captured on Mar 1, 2004 and aggregated over 5 min (left column) with artificially generated TMs using the Gravity Model (right column).Each heatmap represents a single TM in Abilene which comprises 12 nodes.

6 DEMAND-OBLIVIOUS TP

In this section, we present a novel approach to optimize the wavelength assignment followed by an evaluation of the topologies with the most potential regarding reconfigurability.

Motivation. Network (capacity) design is well studied in the context of datacenters, where it is a common objective to maximize the bisection bandwidth, in order to facilitate all-to-all throughput, ideally in a non-blocking fashion [11]. One of the most popular designs in this context is the so-called fat-tree [42], which is easily realizable with off-the-shelf hardware [1]. Herein, the leaf nodes are connected by logical tree topology, with increased capacity towards the core, in order to serve all possible connection pairs evenly. However, unlike WANs, datacenter networks are mostly uniformly designed

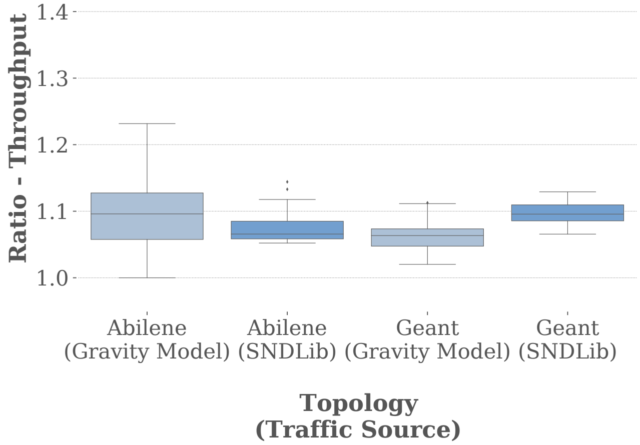


Figure 4: Comparison of synthetic generated traffic (Gravity Model) and real traffic (SNDLib) regarding reconfigurability. A value of 1 means no improvement over uniform TP.

from scratch and operate at relatively small geographical scale, making a direct transition of design ideas infeasible.

Notwithstanding, we can still explore fat-tree designs from the perspective of capacity allocations. Towards the core, the edges have to serve more connection pairs, increasing the capacity. Another way of thinking of these connection pairs is by means of paths that require capacity allocated, where we need more capacity the more paths are assigned to an edge. The latter is a concept that we can directly apply to oblivious WAN capacity designs, attempting to emulate fat-trees in a wide-area setting. To this end, we propose that for all node pairs, we compute suitable connecting paths, and assign capacity relative to the number of paths "living" on these edges. As such, we provide optimized all-to-all connectivity, with the goal of providing efficient demand-oblivious capacity.

We next provide the technical details on our design proposal in §6.1 and then evaluate its performance w.r.t. throughput gains, in comparison to joint TP+TE formulations, in §6.2. We show that our new design proposal can significantly reduce the gap towards joint demand-aware TP+TE formulations to below 5%, on average, and in particular significantly reduces the large gaps still present for uniform capacity designs seen in the last section for a small set of topologies.

6.1 New Demand-Oblivious Capacity Designs

In Algorithm 1, we present a new demand-oblivious approach for maximizing the throughput through optimizing the wavelength assignment.

The idea is to compute all to all shortest paths and determine for each link the number of paths it is included and adjust the wavelength assignment accordingly. We call `DemandObliviousTP(`G, α, β, γ`)` passing the graph G

Algorithm 1 Improved Demand-Oblivious Capacity Design

1: **function** DemandObliviousTP($G = (V, E)$, α, β, γ)
2: $r \leftarrow \{\}$ ▷ The requested capacity on a link
3: $R \leftarrow \{\}$ ▷ The requested capacity on a node
4: $paths \leftarrow$ AllShortestPaths(G)
5: **for** $s \leftarrow 0$ **to** $|V|$ **do**
6: **for** $t \leftarrow s + 1$ **to** $|V|$ **do**
7: **for all** $(i, j) \in paths^{st}$ **do**
8: $r_{ij} \leftarrow r_{ji} \leftarrow r_{ij} + 1$
9: $R_i \leftarrow R_i + 1$
10: $R_j \leftarrow R_j + 1$
11: **end for**
12: **end for**
13: **end for**
14: $w \leftarrow \{\}$ ▷ # Wavelengths assigned to each link
15: **for all** $(i, j) \in E$ **do**
16: $t_i \leftarrow (deg_i \cdot \beta/\alpha) \cdot r_{ij} // R_i$ ▷ // := Integer division
17: $t_j \leftarrow (deg_j \cdot \beta/\alpha) \cdot r_{ij} // R_j$
18: $w_{ij} \leftarrow \min(t_i, t_j, \beta)$
19: **end for**
20: $W \leftarrow \{\}$ ▷ # Wavelengths each node has left to distribute
21: **for** $i \leftarrow 1$ **to** $|V|$ **do**
22: $W_i = deg_i \cdot \beta/\alpha - \sum_{j \in N_G(i)} w_{ij}$ ▷ $N_G(i)$: neighborhood of i
23: **end for**
24: **do**
25: $E' \leftarrow \{(i, j) \in E : W_i > 0$ **and** $W_j > 0$ **and** $w_{ij} < \beta\}$
26: **for all** $(i, j) \in E'$ **do**
27: **if** $W_i > 0$ **and** $W_j > 0$ **and** $w_{ij} < \beta$ **then**
28: $w_{ij} \leftarrow w_{ij} + 1$
29: $W_i \leftarrow W_i - 1$
30: $W_j \leftarrow W_j - 1$
31: **end if**
32: **end for**
33: **while** $E' \neq \emptyset$
34: **return** w
35: **end function**

containing the nodes V and the links E, the node limiter α, the number of wavelengths a fiber link can carry at max β, and the capacity of a single wavelength γ. The variable w stores the assigned wavelengths for each link, W holds the number of wavelengths nodes have left to distribute to its adjacent links. The variable r, the required capacity at each link, is computed by counting all traffic flows traversing the link. For each link (i, j) we evaluate for both i and j separately weighted the number of wavelengths it would assign to the link and assign the minimum of these two values. Finally, in 20-33 we assign the leftover wavelengths evenly to all links. The returned wavelength assignment serves as input for a TE algorithm. For which we used the proposed TE LP in all our evaluations.

6.2 Evaluation of TP Strategies

Fig. 7b shows the improvements by our improved demand-oblivious capacity design from Algorithm 1 on selected topologies. In more detail, we selected the top 10 topologies from the last section with the highest throughput gains, depicting their relative potential over a uniform capacity allocation in Fig. 5. We then plot in Fig. 7b by how much our new designs reduce the gap towards demand-aware optimal joint TP+TE formulations. Herein, a value of 1.0 means exactly

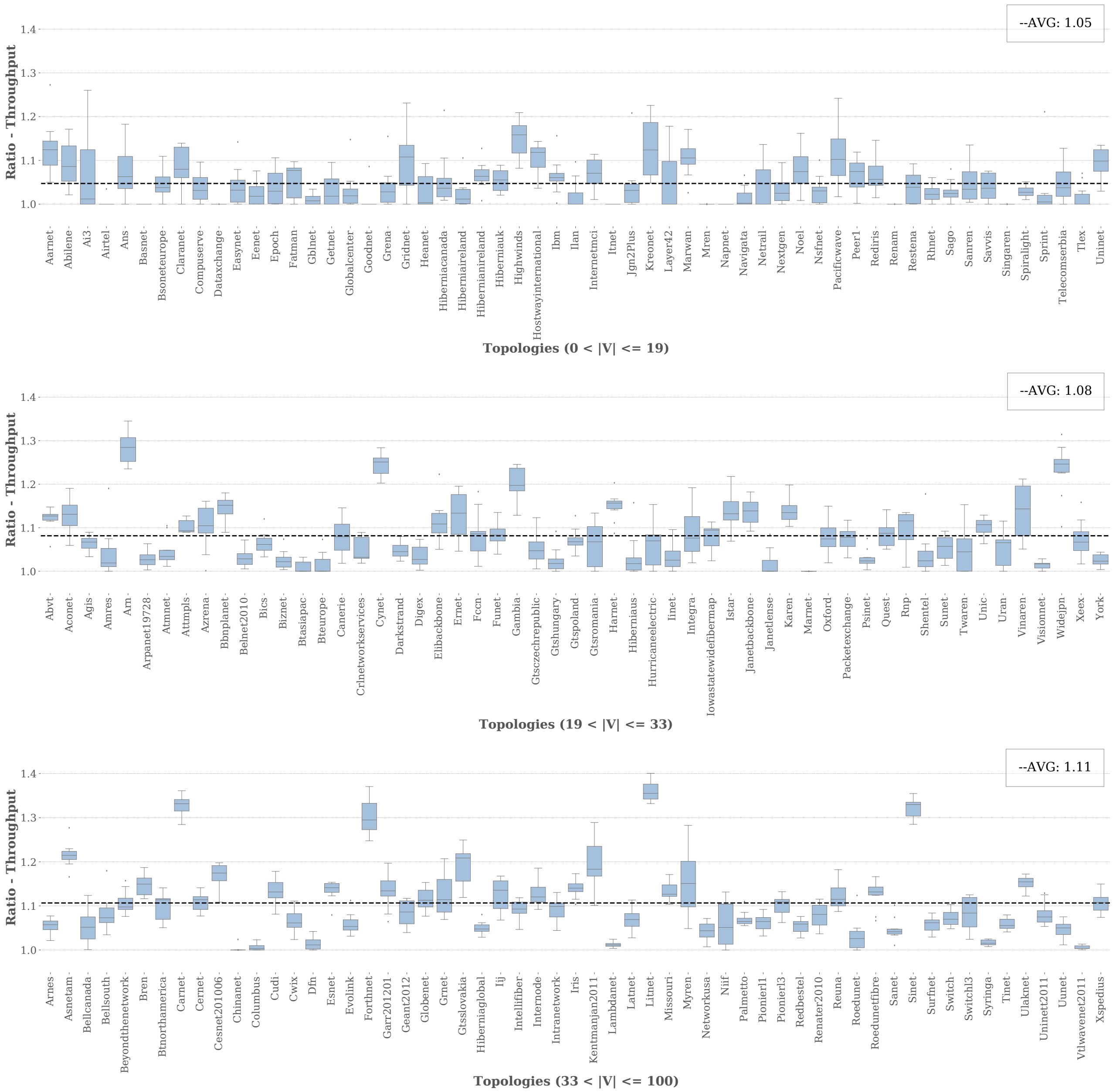
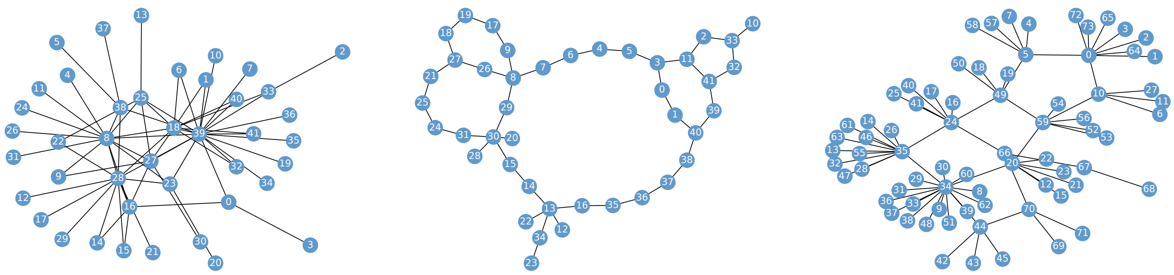


Figure 5: Topology optimization potential for throughput gains by means of topology programming grouped by size $|V|$. Note that for some topologies, the joint TP+TE does not achieve any improvement over uniform TP. Simple star topologies such as Basnet, Itnet, Mren, or the high connectivity in Airtel, Chinanet, Dataxchange, Goodnet, and Napnet prevent optimization.

the same value as joint TP+TE, whereas a value of 0.0 shows no improvement over a uniform capacity allocation.

As we can see, our new demand-oblivious designs are significantly close to optimal demand-aware formulations, closing the gap by up to nearly 90% for Carnet, whereas even for Gambia, the gap reduction is close to 50%. Except for the latter, all gap reductions are over 50%, being close to at least 80% for all of them. The potential gains of throughput optimization by means of TP+TE are again hovering around or below 10%, in comparison to our new oblivious capacity designs.

We note however that our new designs achieve such significant gap reductions only for networks where TP+TE significantly outperforms uniform capacity designs. The lower the potential gains, the lower the gap reduction, and for a good



(a) Chinanet **(b) Lambdanet** **(c) Sinet**

Figure 6: Topological structure of Chinanet, Lambdanet and Sinet from TopologyZoo. The shown node positions are determined using the Kamada-Kawai path-length cost-function.

number of networks, the benefits are close to zero or uniform TP even performs slightly better than our new designs. The latter are topologies where uniform already performs very well or close to an optimal joint TP+TE, such as Chinanet or Lambdanet from Fig. 6.

Notwithstanding, for such cases, one can simply default back to uniform capacity allocations, selecting the best of both demand-oblivious designs depending on the chosen topology. By selecting the better oblivious strategy for each topology (where the capacity allocations stay fixed for each TM), the potential gains of joint TP+TE optimizations are reduced to just 4.76%, in average over all 169 topologies, see Fig. 8.

7 RELATED WORK

While we are not aware of work that investigates oblivious WAN capacity design from a networking perspective, the idea of joint topology programming (TP) and traffic engineering in this context has already been investigated recently. We note that we take a throughput-optimal TP+TE formulation as comparison, whereas the following works aim at faster TP computation times or different objectives.

Jin et al. [36] present Owan, a centralized system to optimize the IP and optical topologies by reconfiguring the network devices, including ROADMs. Their work focuses on optimizing the IP link reconfiguration and routing to minimize the completion time for bulk transfers, with scheduling concerns expanded in [13, 35]. To extend [36] for a multicast setting, Long et al. [46] propose a partially relaxed ILP, where IP and optical layers are jointly optimized in a cross-layer approach, improving deadline scheduling and throughput performance. Although less directly, this work is also related to the work by Gossels et al. [23, 24]. They set their focus on robustness to failures in WAN topologies by utilizing the advantages of ROADMs. In their abstraction model, they consider additionally, regenerators and tails, where a tail refers to a combination of an optical transponder and a router port.

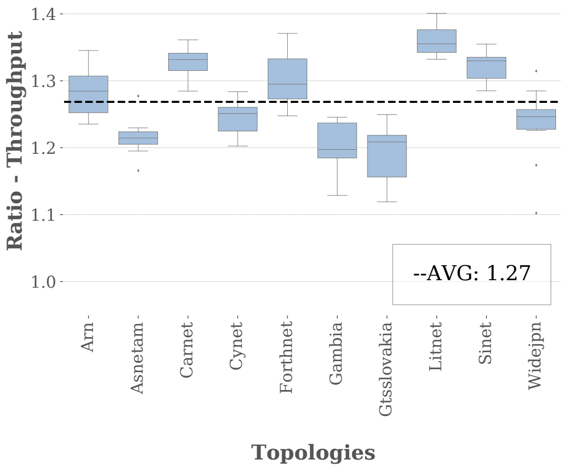
Another line of work is by Foerster et al. [21], who propose an indirect approach of finding the optimal wavelength allocation in optical WAN topologies. To this end, they propose an abstraction, where TP calculations are performed by the TE formulation in a demand-aware fashion.

The achievements in reconfigurable networks in the scope of data centers (DC) [8, 12, 15, 16, 22, 39, 43–45, 49, 59, 83] are significant as well. In DC networks, we can augment the existing topology with reconfigurable links connecting nodes either directly (e.g. free-space optics) or via e.g. optical switches. By jointly solving the TP+TE problems, the efficiency in DC networks can be greatly improved, though the optimization problem remains intractable even for "easy" objective functions [18, 19]. Although inspiring, the underlying assumptions strongly differ compared to those in WANs and a direct adaption of proposed solutions seems impossible, considering the geographical scale, the variety of existing topologies and foremost, the reconfigurability restrictions. However, a related class of reconfigurable networks where WAN designs might potentially benefit from insights is from reconfigurable geometric [7] and satellite networks [5].

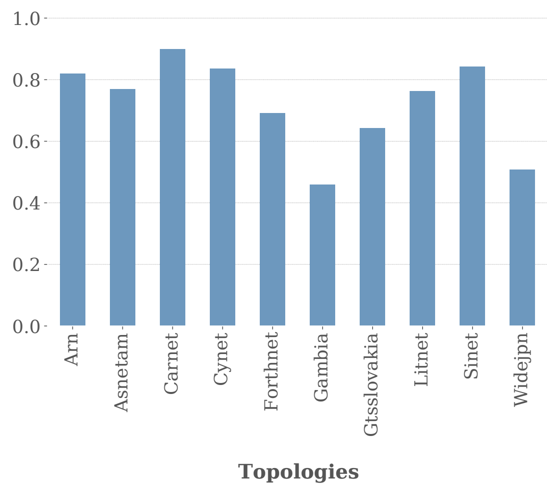
Another class of related work addresses techniques to synthesize traffic matrices. The proprietary nature of Internet traffic matrices led to a particular interest in modeling artificial traffic data [6, 66, 77, 78]. Roughan et al. [66, 78] propose the Gravity Model (GM) to synthesize (but not predict) traffic matrices. To evaluate their approach, they show that for both the real-world data from Abilene and Géant the synthetic traffic fits excellently in respect of their cumulative distribution and complementary cumulative distribution. Tune et al. extend their work by proposing temporal and spatiotemporal models to generate a series of traffic matrices over time.

8 CONCLUSION

Motivated by the popularity of data-centric applications and the emergence of reconfigurable WAN topologies, we investigated the benefits of reconfigurability both under real demands as well as in synthetic scenarios, aiming to identify



(a) Top 10 topologies from Fig. 5 regarding reconfigurability.



(b) Gap reduction by our demand-oblivious capacity design. A value of 1.0 means exactly the same value as joint TP+TE, whereas a value of 0.0 shows no improvement over a uniform capacity allocation.

Figure 7: 7a shows the 10 topologies with the highest potential gain by employing demand-aware TP+TE over our uniform baseline design, whereas 7b shows how much of this potential gain can be achieved by our improved demand-oblivious capacity design proposal. E.g. Widejpn allows for a median improvement of about 25% in throughput over uniform capacity design, but as we close the gap by about 50% with our new strategy, TP+TE only achieves further gains of roughly 12.5%.

under which conditions the performance can be improved the most, by means of a combination of both physical topology programming and traffic engineering.

We found that in most cases already a traffic-independent uniform wavelength deployment yields throughput within 8% of a demand-aware joint topology programming and traffic engineering formulation, while avoiding complex computations and reconfiguration delays. However, we found that in some situations, the benefits can be significant, e.g., allowing up to 40% more throughput over a uniform approach. For these cases, we proposed an oblivious wavelength deployment strategy inspired by state-of-the-art datacenter design, maximizing the all-to-all bandwidth. Our simulations show that our oblivious strategy significantly reduces the gap towards joint topology programming and traffic engineering formulations, while retaining the benefits of a static physical layer.

8.1 Future Work

We understand our work as a first step and hope it can lead to followup work informing the community about the benefits and limitations of more dynamic infrastructures. In particular, it will be interesting to explore refined network models, accounting for reconfiguration and conversion costs, as well as to study the impact of certain graph properties on reconfigurability benefits, such as edge connectivity, treewidth, clustering coefficient, to name only a few examples. It would further be interesting to investigate the trade-off of reconfiguration costs versus benefits [70, 76], in particular in experimental frameworks [80].

Moreover, in this work we chose throughput as an objective function[2], and it is not clear how to extend our efficient oblivious designs to other settings, such as bulk-transfer scheduling [36, 47], CDNs [81] which have scheduling constraints, or [23], which incorporates hardware (placement) choices, e.g., amplifiers, conversions, etc., and failure resilience [27, 82].

Another interesting avenue for future research is to explore improved joint optimizations, including both oblivious algorithms as well as dynamic algorithms reacting to online demand changes, as well as semi-oblivous approaches leveraging long-term traffic predictions.

Lastly, the question of modeling (and provably) optimal demand-aware oblivious designs is still an open research question, for very recent work in this direction we refer to the article by Amir et al. [2].

REPRODUCIBILITY

To facilitate reproducibility, our source code will be made available at https://github.com/tfenz/Programmable-WAN.

ACKNOWLEDGMENTS

We thank our shepherd Lavanya Jose and the reviewers for their valuable feedback. This research received funding from

[2]We refer to Balon et al. [4] for a discussion on TE objective functions.

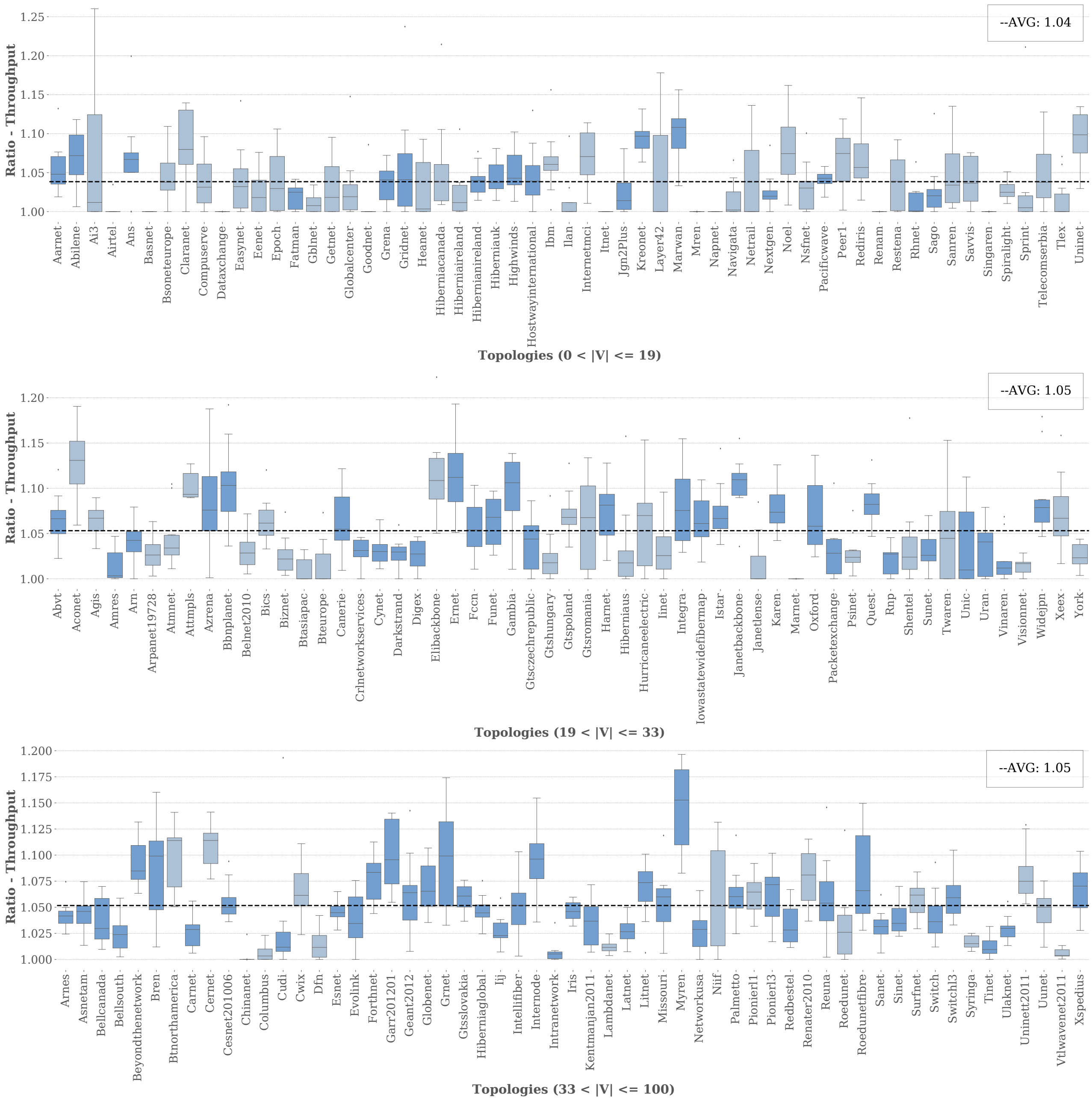


Figure 8: Reduced gap to joint optimization by using the best demand-oblivious topology programming (TP) strategy. (Dark blue boxes signify the topologies where we can reduce the gap to the optimal solution utilizing the Improved Demand-Oblivious TP strategy; Light blue boxes signify that the uniform approach is not improved.)

the European Research Council (ERC) Consolidator project, Self-Adjusting Networks (AdjustNet), grant agreement No. 864228, Horizon 2020, 2020-2025.

REFERENCES

[1] Mohammad Al-Fares, Alexander Loukissas, and Amin Vahdat. 2008. A scalable, commodity data center network architecture. *ACM SIGCOMM computer communication review* 38, 4 (2008), 63–74.

[2] Daniel Amir, Tegan Wilson, Vishal Shrivastav, Hakim Weatherspoon, Robert Kleinberg, and Rachit Agarwal. 2021. Optimal Oblivious Reconfigurable Networks. *CoRR* abs/2111.08780 (2021).

[3] At Long Last – High Performance Contentionless ROADMs. Last Accessed: 2021-09-14. https://ribboncommunications.com/company/media-center/blog/long-last-high-performance-contentionless-roadms.

[4] Simon Balon, Fabian Skivée, and Guy Leduc. 2006. How Well Do Traffic Engineering Objective Functions Meet TE Requirements?. In *Networking (Lecture Notes in Computer Science, Vol. 3976)*. Springer, 75–86.

[5] Debopam Bhattacherjee and Ankit Singla. 2019. Network topology design at 27, 000 km/hour. In *CoNEXT*. ACM, 341–354.

[6] Rhys Alistair Bowden, Matthew Roughan, and Nigel G. Bean. 2014. COLD: PoP-level Network Topology Synthesis. In *CoNEXT*. ACM, 173–184.

[7] Esra Ceylan, Klaus-Tycho Foerster, Stefan Schmid, and Katsiaryna Zaitsava. 2021. Demand-Aware Plane Spanners of Bounded Degree. In *Networking*. IEEE, 1–9.

[8] Li Chen, Kai Chen, Zhonghua Zhu, Minlan Yu, George Porter, Chunming Qiao, and Shan Zhong. 2017. Enabling wide-spread communications on optical fabric with megaswitch. In *NSDI*. USENIX Association, 577–593.

[9] Angela L Chiu et al. 2011. Architectures and protocols for capacity efficient, highly dynamic and highly resilient core networks. *IEEE/OSA Journal of Optical Communications and Networking* 4, 1 (2011), 1–14.

[10] Cisco. 2020. Cisco Annual Internet Report (2018–2023) White Paper.

[11] Charles Clos. 1953. A study of non-blocking switching networks. *Bell System Technical Journal* 32, 2 (1953), 406–424.

[12] Wenkai Dai, Klaus-Tycho Foerster, David Alexander Fuchssteiner, and Stefan Schmid. 2020. Load-Optimization in Reconfigurable Networks: Algorithms and Complexity of Flow Routing. In *38th International Symposium on Computer Performance, Modeling, Measurements and Evaluation 2020*. http://eprints.cs.univie.ac.at/6435/

[13] Michael Dinitz and Benjamin Moseley. 2020. Scheduling for Weighted Flow and Completion Times in Reconfigurable Networks. In *INFOCOM*. IEEE, 1043–1052.

[14] Efficiently Supporting Aggressive Network Capacity Growth in Next-Generation ROADM Networks. Last Accessed: 2021-09-14. https://resource.lumentum.com/s3fs-public/technical-library-items/next-genroadm-wp-oc-ae.pdf.

[15] Nathan Farrington, George Porter, Sivasankar Radhakrishnan, Hamid Hajabdolali Bazzaz, Vikram Subramanya, Yeshaiahu Fainman, George Papen, and Amin Vahdat. 2010. Helios: a hybrid electrical/optical switch architecture for modular data centers. In *SIGCOMM*. ACM, 339–350.

[16] Thomas Fenz, Klaus-Tycho Foerster, Stefan Schmid, and Anaïs Villedieu. 2020. Efficient non-segregated routing for reconfigurable demand-aware networks. *Computer Communications* 164 (2020), 138 – 147.

[17] Mark Filer, Jamie Gaudette, Monia Ghobadi, Ratul Mahajan, Tom Issenhuth, Buddy Klinkers, and Jeff Cox. 2016. Elastic optical networking in the Microsoft cloud. *IEEE/OSA Journal of Optical Communications and Networking* 8, 7 (2016), A45–A54.

[18] Klaus-Tycho Foerster, Manya Ghobadi, and Stefan Schmid. 2018. Characterizing the algorithmic complexity of reconfigurable data center architectures. In *ANCS*. ACM, 89–96.

[19] Klaus-Tycho Foerster, Maciej Pacut, and Stefan Schmid. 2019. On the Complexity of Non-Segregated Routing in Reconfigurable Data Center Architectures. *Comput. Commun. Rev.* 49, 2 (2019), 2–8.

[20] Klaus-Tycho Foerster and Stefan Schmid. 2019. Survey of Reconfigurable Data Center Networks: Enablers, Algorithms, Complexity. *SIGACT News* 50, 2 (2019), 62–79.

[21] Klaus-Tycho Foerster, Long Luo, and Manya Ghobadi. 2020. OptFlow: A Flow-Based Abstraction for Programmable Topologies. In *Proceedings of the Symposium on SDN Research* (San Jose, CA, USA) *(SOSR '20)*. Association for Computing Machinery, New York, NY, USA, 96–102.

[22] Monia Ghobadi, Ratul Mahajan, Amar Phanishayee, Nikhil R. Devanur, Janardhan Kulkarni, Gireeja Ranade, Pierre-Alexandre Blanche, Houman Rastegarfar, Madeleine Glick, and Daniel C. Kilper. 2016. ProjecToR: Agile Reconfigurable Data Center Interconnect. In *SIGCOMM*. ACM, 216–229.

[23] Jennifer Gossels. 2020. *Joint Optimization for Robust Network Design and Operation*. Ph. D. Dissertation. Princeton University.

[24] Jennifer Gossels, Gagan Choudhury, and Jennifer Rexford. 2019. Robust network design for IP/optical backbones. *J. Opt. Commun. Netw.* 11, 8 (Aug 2019), 478–490.

[25] Chen Griner, Johannes Zerwas, Andreas Blenk, Stefan Schmid, Manya Ghobadi, and Chen Avin. 2022. Cerberus: The Power of Choices in Datacenter Topology Design (A Throughput Perspective). In *Proc. ACM SIGMETRICS*.

[26] Gurobi Optimizer 9.0.1. Last Accessed: 2021-09-14. www.gurobi.com/.

[27] Matthew Nance Hall, Paul Barford, Klaus-Tycho Foerster, Manya Ghobadi, William Jensen, and Ramakrishnan Durairajan. 2021. Are WANs Ready for Optical Topology Programming?. In *OptSys@SIGCOMM*. ACM, 28–33.

[28] Matthew Nance Hall, Klaus-Tycho Foerster, Stefan Schmid, and Ramakrishnan Durairajan. 2021. A Survey of Reconfigurable Optical Networks. *Opt. Switch. Netw.* 41 (2021), 100621.

[29] Tad Hofmeister, Vijay Vusirikala, and Bikash Koley. 2016. How can flexibility on the line side best be exploited on the client side?. In *Optical Fiber Communication Conference*. Optical Society of America, W4G–4.

[30] Chi-Yao Hong et al. 2018. B4 and after: managing hierarchy, partitioning, and asymmetry for availability and scale in google's software-defined WAN. In *SIGCOMM*. ACM, 74–87.

[31] Chi-Yao Hong, Srikanth Kandula, Ratul Mahajan, Ming Zhang, Vijay Gill, Mohan Nanduri, and Roger Wattenhofer. 2013. Achieving high utilization with software-driven WAN. In *SIGCOMM*. ACM, 15–26.

[32] Mingfeng Huang, Anfeng Liu, Neal N Xiong, Tian Wang, and Athanasios V Vasilakos. 2020. An effective service-oriented networking management architecture for 5G-enabled internet of things. *Computer Networks* (2020), 107208.

[33] Yishen Huang, Craig L Gutterman, Payman Samadi, Patricia B Cho, Wiem Samoud, Cédric Ware, Mounia Lourdiane, Gil Zussman, and Keren Bergman. 2017. Dynamic mitigation of EDFA power excursions with machine learning. *Optics express* 25, 3 (2017), 2245–2258.

[34] Sushant Jain, Alok Kumar, Subhasree Mandal, Joon Ong, Leon Poutievski, Arjun Singh, Subbaiah Venkata, Jim Wanderer, Junlan Zhou, Min Zhu, et al. 2013. B4: Experience with a globally-deployed software defined WAN. *ACM SIGCOMM Computer Communication Review* 43, 4 (2013), 3–14.

[35] Su Jia, Xin Jin, Golnaz Ghasemiesfeh, Jiaxin Ding, and Jie Gao. 2017. Competitive analysis for online scheduling in software-defined optical WAN. In *INFOCOM*. IEEE, 1–9.

[36] Xin Jin, Yiran Li, Da Wei, Siming Li, Jie Gao, Lei Xu, Guangzhi Li, Wei Xu, and Jennifer Rexford. 2016. Optimizing Bulk Transfers with Software-Defined Optical WAN. In *SIGCOMM*. ACM, 87–100.

[37] Tomihisa Kamada, Satoru Kawai, et al. 1989. An algorithm for drawing general undirected graphs. *Information processing letters* 31, 1 (1989), 7–15.

[38] Srikanth Kandula, Ishai Menache, Roy Schwartz, and Spandana Raj Babbula. 2014. Calendaring for wide area networks. In *SIGCOMM*. ACM, 515–526.

[39] Srikanth Kandula, Jitendra Padhye, and Paramvir Bahl. 2009. Flyways To De-Congest Data Center Networks. In *HotNets*.

[40] S. Knight, H.X. Nguyen, N. Falkner, R. Bowden, and M. Roughan. 2011. The Internet Topology Zoo. *Selected Areas in Communications, IEEE Journal on* 29, 9 (october 2011), 1765 –1775. https://doi.org/10.1109/JSAC.2011.111002

[41] Praveen Kumar, Yang Yuan, Chris Yu, Nate Foster, Robert Kleinberg, Petr Lapukhov, Chiunlin Lim, and Robert Soulé. 2018. Semi-Oblivious Traffic Engineering: The Road Not Taken. In *NSDI*. USENIX Association, 157–170.

[42] Charles E Leiserson. 1985. Fat-trees: universal networks for hardware-efficient supercomputing. *IEEE transactions on Computers* 100, 10

(1985), 892–901.

[43] He Liu, Feng Lu, Alex Forencich, Rishi Kapoor, Malveeka Tewari, Geoffrey M Voelker, George Papen, Alex C Snoeren, and George Porter. 2014. Circuit switching under the radar with REACToR. In *NSDI*. USENIX Association, 1–15.

[44] He Liu, Matthew K. Mukerjee, Conglong Li, Nicolas Feltman, George Papen, Stefan Savage, Srinivasan Seshan, Geoffrey M. Voelker, David G. Andersen, Michael Kaminsky, George Porter, and Alex C. Snoeren. 2015. Scheduling techniques for hybrid circuit/packet networks. In *CoNEXT*. ACM, 41:1–41:13.

[45] Yunpeng James Liu, Peter Xiang Gao, Bernard Wong, and Srinivasan Keshav. 2014. Quartz: a new design element for low-latency DCNs. *ACM SIGCOMM Computer Communication Review* 44, 4 (2014), 283–294.

[46] Long Luo, Klaus-Tycho Foerster, Stefan Schmid, and Hongfang Yu. 2020. Deadline-Aware Multicast Transfers in Software-Defined Optical Wide-Area Networks. *IEEE J. Sel. Areas Commun.* 38, 7 (2020), 1584–1599.

[47] Long Luo, Hongfang Yu, Klaus-Tycho Foerster, Max Noormohammadpour, and Stefan Schmid. 2020. Inter-Datacenter Bulk Transfers: Trends and Challenges. *IEEE Netw.* 34, 5 (2020), 240–246.

[48] Hannah Ritchie Max Roser and Esteban Ortiz-Ospina. 2015. Internet. *Our World in Data* (2015). https://ourworldindata.org/internet.

[49] William M. Mellette, Rob McGuinness, Arjun Roy, Alex Forencich, George Papen, Alex C. Snoeren, and George Porter. 2017. RotorNet: A Scalable, Low-complexity, Optical Datacenter Network. In *SIGCOMM*. ACM, 267–280.

[50] Ali Nauman, Yazdan Ahmad Qadri, Muhammad Amjad, Yousaf Bin Zikria, Muhammad Khalil Afzal, and Sung Won Kim. 2020. Multimedia Internet of Things: A comprehensive survey. *IEEE Access* 8 (2020), 8202–8250.

[51] NetworkX 2.4. Last Accessed: 2021-09-14. https://networkx.org/documentation/networkx-2.4/.

[52] NumPy 1.18.1. Last Accessed: 2021-09-14. https://numpy.org/devdocs/release/1.18.1-notes.html.

[53] Shoichiro Oda, Masatake Miyabe, Setsuo Yoshida, Toru Katagiri, Yasuhiko Aoki, Jens C Rasmussen, Martin Birk, and Kathy Tse. 2016. Demonstration of an autonomous, software controlled living optical network that eliminates the need for pre-planning. In *Optical Fiber Communication Conference*. Optical Society of America, W2A–44.

[54] Open ROADM Multi-Source Agreement (MSA). Last Accessed: 2021-09-14. http://openroadm.org/.

[55] S. Orlowski, M. Pióro, A. Tomaszewski, and R. Wessäly. 2007. SNDlib 1.0–Survivable Network Design Library. In *Proceedings of the 3rd International Network Optimization Conference (INOC 2007), Spa, Belgium*. http://sndlib.zib.de, extended version accepted in Networks, 2009..

[56] Sebastian Orlowski, Roland Wessäly, Michal Pióro, and Artur Tomaszewski. 2010. SNDlib 1.0—Survivable network design library. *Networks: An International Journal* 55, 3 (2010), 276–286.

[57] Maciej Pacut, Wenkai Dai, Alexandre Labbe, Klaus-T. Foerster, and Stefan Schmid. 2021. Improved scalability of demand-aware datacenter topologies with minimal route lengths and congestion. *Performance Evaluation* 152 (2021), 102238.

[58] Sterling Perrin. 2015. Next-generation ROADM architectures and benefits. *Heavy Reading White Paper, Mar* (2015).

[59] George Porter, Richard Strong, Nathan Farrington, Alex Forencich, Pang Chen-Sun, Tajana Rosing, Yeshaiahu Fainman, George Papen, and Amin Vahdat. 2013. Integrating microsecond circuit switching into the data center. *ACM SIGCOMM Computer Communication Review* 43, 4 (2013), 447–458.

[60] Renfrey B Potts and Robert M Oliver. 1972. Flows in Transportation Networks.

[61] Pentti Pöyhönen. 1963. A tentative model for the volume of trade between countries. *Weltwirtschaftliches Archiv* (1963), 93–100.

[62] Qifan Pu, Ganesh Ananthanarayanan, Peter Bodik, Srikanth Kandula, Aditya Akella, Paramvir Bahl, and Ion Stoica. 2015. Low latency geo-distributed data analytics. *ACM SIGCOMM Computer Communication Review* 45, 4 (2015), 421–434.

[63] Python 3.7.6. Last Accessed: 2021-09-14. www.python.org/downloads/release/python-376/.

[64] Harald Räcke. 2008. Optimal hierarchical decompositions for congestion minimization in networks. In *Proceedings of the fortieth annual ACM symposium on Theory of computing*. 255–264.

[65] ROADM. Last Accessed: 2021-09-14. https://www.adva.com/en/products/technology/roadm.

[66] Matthew Roughan. 2005. Simplifying the synthesis of Internet traffic matrices. *ACM SIGCOMM Computer Communication Review* 35, 5 (2005), 93–96.

[67] Akio Sahara, Yukio Tsukishima, Tetsuo Takahashi, Youhei Okubo, Kazuhisa Yamada, Kazuhiro Matsuda, and Atsushi Takada. 2009. Demonstration of colorless and directed/directionless ROADMs in router network. In *2009 Conference on Optical Fiber Communication*. IEEE, 1–3.

[68] SciPy 1.1.0. Last Accessed: 2021-09-14. https://docs.scipy.org/doc/scipy/reference/release.1.1.0.html.

[69] Niklas Semmler, Matthias Rost, Georgios Smaragdakis, and Anja Feldmann. 2020. Edge replication strategies for wide-area distributed processing. In *EdgeSys@EuroSys*. ACM, 1–6.

[70] Rachee Singh, Monia Ghobadi, Klaus-Tycho Foerster, Mark Filer, and Phillipa Gill. 2017. Run, Walk, Crawl: Towards Dynamic Link Capacities. In *HotNets*. ACM, 143–149.

[71] Rachee Singh, Manya Ghobadi, Klaus-Tycho Foerster, Mark Filer, and Phillipa Gill. 2018. RADWAN: rate adaptive wide area network. In *SIGCOMM*. ACM, 547–560.

[72] SndLib. Last Accessed: 2021-09-14. http://sndlib.zib.de/home.action?show=/docu.formats.gml.action.

[73] Zhongda Tian. 2020. Chaotic characteristic analysis of network traffic time series at different time scales. *Chaos, Solitons & Fractals* 130 (2020), 109412.

[74] Jan Tinbergen. 1962. *Suggestions for an international economic policy*. Periodicals Service Company.

[75] Topology Zoo. Last Accessed: 2021-09-14. http://www.topology-zoo.org/dataset.html.

[76] Shih-Hao Tseng. 2019. Perseverance-Aware Traffic Engineering in Rate-Adaptive Networks with Reconfiguration Delay. In *ICNP*. IEEE, 1–10.

[77] Paul Tune and Matthew Roughan. 2015. Spatiotemporal Traffic Matrix Synthesis. In *SIGCOMM*. ACM, 579–592.

[78] Paul Tune, Matthew Roughan, H Haddadi, and O Bonaventure. 2013. Internet traffic matrices: A primer. *Recent Advances in Networking* 1 (2013), 1–56.

[79] International Telecommunication Union. Last Accessed: 2021-09-14. Individuals Using the Internet. https://www.itu.int/en/ITU-D/Statistics/Pages/stat/default.aspx.

[80] Johannes Zerwas, Chen Avin, Stefan Schmid, and Andreas Blenk. 2021. ExRec: Experimental Framework for Reconfigurable Networks Based on Off-the-Shelf Hardware. In *ACM/IEEE Symposium on Architectures for Networking and Communications Systems (ANCS)*.

[81] Johannes Zerwas, Ingmar Poese, Stefan Schmid, and Andreas Blenk. 2021. On the Benefits of Joint Optimization of Reconfigurable CDN-ISP Infrastructure. *IEEE Transactions on Network and Service Management* (2021). to appear.

[82] Zhizhen Zhong, Manya Ghobadi, Alaa Khaddaj, Jonathan Leach, Yiting Xia, and Ying Zhang. 2021. ARROW: restoration-aware traffic engineering. In *SIGCOMM*. ACM, 560–579.

[83] Xia Zhou, Zengbin Zhang, Yibo Zhu, Yubo Li, Saipriya Kumar, Amin Vahdat, Ben Y Zhao, and Haitao Zheng. 2012. Mirror mirror on the ceiling: Flexible wireless links for data centers. *ACM SIGCOMM Computer Communication Review* 42, 4 (2012), 443–454.