

On the Benefits of Joint Optimization of Reconfigurable CDN-ISP Infrastructure

Johannes Zerwas, Ingmar Poesse, Stefan Schmid, and Andreas Blenk

Abstract—ISP networks have become a critical infrastructure in our society. Traffic in these networks is growing and is increasingly dominated by a small number of large CDNs connecting at multiple locations. Simultaneously, the networks are becoming more flexible, in terms of routing, CDN user mapping, and also regarding the IP topology: emerging optical technologies allow to flexibly reconfigure the network.

This paper studies the potential gains of these reconfiguration flexibilities. The idea is to make CDN-ISP infrastructure demand-aware, that is, to re-optimize it towards the changing end-user demands over time. We present an optimization framework and conduct an extensive evaluation using data from a large European ISP. We find that such a reconfigurable infrastructure has indeed a high potential: by leveraging spatial and diurnal traffic patterns, the efficiency of ISP networks and CDNs is improved significantly. Specifically, the required backbone capacity is reduced by 15% while reducing path lengths by 30%, on average and during the critical peak hour. Moreover, such infrastructures can leverage re-optimizations during specific events, like the COVID-19 pandemic, and under link failures. We optimistically assume a cooperative environment of ISPs and CDNs, and we conclude by discussing trends that foster the identified benefits in practice.

Index Terms—Optical networks, ISP, Hyper-giants, CDNs, Mixed Integer Programming, Network planning, Reconfiguration

I. INTRODUCTION

Communication networks in general and ISP networks in particular form a critical backbone of the digital society. Given the popularity of data-centric applications, related to health, science, social networking, business and entertainment, it is expected that the traffic carried by these networks will continue to grow explosively, especially to and from datacenters [1]. The COVID-19 pandemic has further highlighted the need for an efficient and reliable communication infrastructure, which is now also critical for, e.g., online teaching, virtual conferences and health services. As traffic demands continue to rise steeply, state-of-the-art approaches to design and operate networks may become expensive and inefficient.

In order to serve traffic workloads efficiently, over the last years, researchers alongside service providers have made

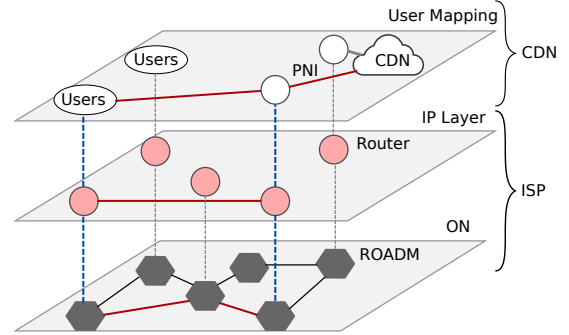


Figure 1: ISPs operate a reconfigurable Optical Network (ON) to connect their customers to content providers such as CDNs. Large CDNs can connect at multiple peering points, which allows the ISP to optimize traffic steering along three dimensions: ON topology, IP layer, and peering point selection.

a concerted effort to innovate at several layers of the networking stack. The proposals render networks more flexible and *demand-aware* and allow to exploit the specific spatio-temporal structure in the demand. For instance, ad-hoc traffic engineering in wide-area networks (WANs) has been replaced with software-defined centralized controllers (e.g., Google B4 [2] and Microsoft SWAN [3]), commodity hardware load-balancers have been replaced with software load-balancers [4], [5], and switch vendors' management APIs have been replaced with in-house switch stacks [6], [7]. With each such technology investment, providers are improving the performance and cost-of-ownership of networks by adding reconfigurability to individual components.

The next frontier toward more adaptive networks is the *physical (optical) layer*: emerging optical technologies allow to reconfigure the network topology in a demand-aware manner [8], [9] within hours or even minutes [10]–[12], e.g., using WDM-based technologies, ROADMs, and elastic optical networks. This in principle allows to further improve the efficiency of networks: by providing “shortcuts” between more frequently communicating sites, the overall traffic may be reduced even in the short term, saving resources and improving latency [11], [13]–[15].

However, only little is known today about the potential benefits (e.g., related to performance, energy consumption, CapEx, QoS) and limitations of more adaptive optical networks, *jointly* optimizing along the different dimensions of flexibility, and hence also accounting for topological reconfigurability.

In this paper, we are particularly interested in the potential benefits of using adaptive and jointly optimized networks to

Johannes Zerwas is with the Chair of Communication Networks, Department of Electrical and Computer Engineering, Technical University of Munich, Munich, Germany (e-mail: johannes.zerwas@tum.de).

Ingmar Poesse is with BENOCS GmbH, Berlin, Germany.

Stefan Schmid is with TU Berlin, Berlin, Germany and Fraunhofer SIT, Berlin, Germany.

Andreas Blenk is with the Chair of Communication Networks, Department of Electrical and Computer Engineering, Technical University of Munich, Munich, Germany and with the Communication Technologies Group, Faculty of Computer Science, University of Vienna, Vienna, Austria.

serve CDN traffic: the traffic of a small number of so-called *hyper-giants* [16] constitutes the majority of the ISP's workload today. For instance, the top-10 *hyper-giants* can make up to 75% of the total traffic [1]. This traffic is also one of the main reasons for ISPs to upgrade their infrastructures [16]. The study of CDN traffic is also interesting since, in addition to its enormous volume, *hyper-giants* increasingly interconnect with eyeball networks through *multiple* locations, which introduces an optimization opportunity to handle this traffic efficiently.

In summary, we consider the potential efficiency and performance benefits of reconfigurable combined CDN-ISP infrastructure. Fig. 1 illustrates the scenario: the ISP operates a reconfigurable Optical Network (ON) to connect its customers to a *hyper-giant* peering at multiple locations. Since large CDNs connect at multiple peering points, in principle, operators can optimize the *hyper-giants'* traffic steering on three fronts: on the optical network topology, on the IP layer, and in terms of peering point selection (i.e., mapping end-users to the "best" ingress point). Assuming a cooperative environment of CDNs and ISPs, we evaluate how CDN peering point selection can jointly be optimized with the ISP network and adapted towards changing end-user demands over time.

A. Our Contributions

This paper analyzes the benefits of exploiting the reconfiguration flexibilities available in modern network infrastructure, for an adaptive and demand-aware re-optimization along three fronts, routing, CDN user mapping, and the IP topology. Optimization of each of these layers or of combinations of two layers have already been evaluated in prior works. In contrast, we study to what extent and how a reconfigurable optical network topology can be combined with a clever request mapping to optimize *hyper-giant* traffic routing in an ISP network. We present an optimization framework which, given the peering locations (PoPs) of *hyper-giants* as well as the end-users' demands, *jointly* optimizes and adapts the mapping from end-users to the *hyper-giants'* PoPs, the IP layer topology and routing, as well as the routing in the optical domain.

We evaluate our approach empirically based on measurement data. We find that our approach can significantly lower the ISP's network loads during the critical peak hour but also on average, and reduce the required backbone capacity by up to 15%. We also provide insights into the required amount and frequency of re-optimizations (the above major benefits are obtained using moderate hourly reconfigurations), the predictability of required optimization (optimizations repeat fairly well on a daily basis and can, hence, be planned ahead of time), as well as the benefits of adaptive reconfigurations during events such as the COVID-19 pandemic (2019-X) [17] or under link failures.

Our analysis of the potential benefits of dynamic re-optimizations is optimistic in that it assumes a cooperative environment; hence, we conclude by discussing deployment scenarios to exploit these benefits in practice.

Reproducibility and research artifacts. While we are not allowed to share the measured ISP traffic patterns due to privacy concerns, in order to facilitate followup work, we will

make our framework implementation publicly available.¹ We will also contribute all our evaluation results to the research community.

B. Organization

The remainder of this paper is organized as follows: §II describes challenges that come from *hyper-giant* dominated workloads and elaborates on the opportunities in today's increasingly flexible networks. §III introduces the optimization framework. The performance of our approach in comparison to existing approaches is evaluated using real data from a large European ISP in §IV. §V elaborates trends that support deployment of our approach in CDNs' and ISPs' infrastructures. Finally, we provide a brief review of related works (§VI) and conclude our contribution (§VII).

II. BACKGROUND AND MOTIVATION

To understand the challenges and optimization opportunities for ISPs, let us revisit the typical architecture of today's ISP networks in more details. To connect their user- or other networks-facing (edge) routers at different locations, ISPs operate optical networks (ONs) that are usually shared based on Dense Wavelength Division Multiplexing (DWDM). DWDM systems consist of reconfigurable optical add/drop multiplexers (ROADMs) that provide connectivity between two nodes in the optical topology, by switching/routing lightpaths through the fibers. IP routers connect to the ONs via the ROADMs. For every lightpath that connects two IP routers, a transceiver/port must be available at both ends of the path. A single lightpath provides a fixed capacity. Multiple lightpaths can be aggregated to provide a single IP link with the accumulated capacity of the single lightpaths.

Combining multiple IP links, the ISP can build a topology where traffic that enters or leaves the network through the aggregation network towards end-users, or via peerings or internet exchanges towards other networks, is routed.

A. The Challenge: Hyper-giant Traffic

Given their size and the important role *hyper-giants* play today, the efficient delivery of their end-user traffic has become a main concern of ISPs [18]. An efficient delivery, i.e., providing a good service while keeping network loads low, is also in the interest of the *hyper-giants*, which aim to offer a low latency and high QoE to their customers.

Today, *hyper-giants* and ISPs often leverage direct connections between their autonomous systems (so-called private network interconnects, short PNI) which gives them full control over the links [19]. In fact, to inject traffic into the ISPs network closer to the end-users, ISPs and *hyper-giants* typically connect at *multiple* geographic locations [19], [20]. This introduces steerability of the ingress point which distinguishes *hyper-giant* traffic from other traffic, henceforth called background traffic. Although background traffic can be much more volatile and exhibit more dense connectivity, the

¹<https://github.com/tum-lkn/hypergiant-isp-optimization>

large volume of *hyper-giant* traffic often dominates in ISP networks.

Reconfiguration of ISPs' topology and routing to accommodate diurnal patterns of background traffic has been shown to be feasible [11]. The diversity of ingress points, however, introduces the challenge of how to map end-users to the "best" ingress point and adds a new dimension to the problem [16], [21], [22]. The volume and spatial-distribution of demands from end-users significantly vary over time, not only due to regular diurnal patterns, but also as a result of large events, e.g., [23]–[25] or the COVID-19 pandemic [26]. While a fixed user-to-ingress point mapping may work well at one point in time, it can lead to congested peerings or network paths at a different time. On the one hand, due to its sheer volume, a sub-optimal topology and routing for *hyper-giants'* traffic results in large overheads for ISPs, e.g., in terms of resource efficiency. On the other hand, the deployed IP topology of the ISP affects shortest paths between end-users and the *hyper-giants*. Hence, end-user mappings, the ISP's IP topology, and the routing interact with each other, which motivates a joint optimization approach accounting for reconfiguration.

B. Opportunity: Demand-Aware Re-Optimization

Re-optimizations accounting for the changing demand can be beneficial at different time scales. A first opportunity concerns diurnal traffic patterns in large eyeball networks. Typical traffic patterns show a large difference between the total traffic volume during the peak hour (typically in the evening) and the hour with the smallest amount of traffic (typically in the night). Our analysis of the traffic volumes arising at a large European ISP reveals that the difference can be as large as factor of 7. An interesting use case for more adaptive networks hence regards the joint re-optimization of CDN-user-to-ingress-point mapping and the ISP's IP topology. Re-optimizing the topology frees up capacities on the fibers, enabling the ISP to offer additional, time-of-day-based services. One option is to temporally sell bare lightpaths to customers that want to directly connect their sites during the low traffic hours (λ -service), e.g., for synchronization. Additional benefits may arise in terms of energy consumption: transceivers may be switched off, and potentially also IP nodes [27].

A second opportunity for potential savings for the ISP and better performance for end-users stems from joint re-optimization on larger scales, e.g., monthly. To accommodate growing demands, both *hyper-giants* and ISPs continuously invest into server, network, and peering infrastructure. Besides a simple upgrade of the capacity at already existing points of presence (PoP), this includes also interconnecting at new locations. While *hyper-giants'* mapping systems usually try to leverage these new ingress points, joint re-optimization including topology and routing introduces also cost savings for the ISP, e.g., by using available transceivers and fibers more efficiently.

C. Enablers: Operational Flexibilities

The proposed joint reconfiguration is fostered by four trends arising in ISP networks that provide operational flexibility:

1. **Flexible IP topology.** The operation and deployment of lightpaths induces costs for the ISP (e.g., CAPEX for the required transceivers or energy costs for the operation of the lightpaths). Hence, ISPs typically aim at deploying as few lightpaths as possible while serving all demands, and accounting for additional requirements, e.g., related to resilience or business policies. Changing the IP topology on a long time-scale, e.g., for maintenance or to accommodate organic growth of traffic, is already common operation. Recent developments in optical switching and advances in Software-Defined Networks (SDN) in the optical domain render also short-term (re-)deployment of lightpaths feasible [28].

However, the time required for changing IP links is still in the order of minutes as multiple steps are necessary: First, the correct paths, including wavelengths, in the optical domain have to be set up. Whereas setting the configuration on a single ROADM is doable within a few seconds [12], setting an entire path can take several minutes [10], [11]. Afterwards, interfaces on the IP routers have to be set up and eventually, updated link characteristics are communicated to the routing entity, e.g., the Path Computation Element (PCE) [29]. Overall, this process can happen on the scale of minutes [11] and results in a first optimization opportunity: a programmable network controller can be used to reconfigure the IP topology throughout the day to optimize it for the changing demands.

2. **Flexible traffic engineering.** On top of the established IP topology, ISPs typically perform traffic engineering using IP or MPLS routing to avoid congestion in the backbone [30]. Recently, novel source-based routing approaches emerged in the context of Segment Routing (SR). Here, each node and adjacency is assigned a unique label (Segment ID). The edge routers push a stack of Segment IDs to forwarded packets which are then used to successively forward the packet [31]. Deployment of SR policies to the edge routers is often done using PCE [29]. This allows adaptable IP routing without propagating updates through the network and results in re-configuration times in the order of seconds [32]. In particular, when Adjacency Segment IDs are used to describe the path, re-convergence of the IGP due to IP topology changes is not necessary. Reachability information or shortest paths are not maintained on the routers in this case. The central controller configures Segment IDs for added or removed links and can update rules on the edge routers accordingly.

3. **Flexible user mapping.** Many *hyper-giants* employ clever schemes to map end-user requests to the desired ingress point/server, e.g., [33]–[36], often using DNS. The *hyper-giant's* DNS system returns different IP addresses to end-users to route them via a specific peering and thereby, to optimize, e.g., their latency or the load of its servers. The benefits of flexible mappings that automatically adapt with the state of the network, have been shown to significantly reduce latency for end-users as well as backbone traffic load for ISPs, especially in CDN-ISP collaboration systems such as PaDIS [21] or FlowDirector [16].

Reconfiguration times depend not only on the agreed channel between *hyper-giant* and ISP, e.g., fully automated using ALTO [37], but also on the time until DNS changes are propagated to the end-users. Recent analyses show that this

can be in the order of minutes [38]. This control over the ingress point hence describes our third enabler for short-term reconfigurability. Throughout this paper, we use the terms *PoP assignment* and *user mapping* interchangeably.

4. Centralized control for joint optimization. While the enablers discussed above are promising, the highest benefit can be obtained by *joint* optimization along all three interacting layers, as we elaborate in the evaluation. For example, changing the PoP assignment without sufficient capacity on the network path may lead to congestion. The SDN paradigm which is also emerging in ISP networks [11], is the fourth enabler: with its centralized control, this problem can be overcome.

III. JOINT OPTIMIZATION FRAMEWORK

In order to study the potential benefits of re-optimizing reconfigurable networks, we formulate a joint optimization framework. Given the *hyper-giant* and background traffic, our goal is to find assignments of end-users to *hyper-giant* peering locations, the routing on the IP layer, the capacities of the IP links, and their routing in the optical domain. The objective is to minimize the network capacity and account for re-configurations. This is subject to fulfilling all demands without violating capacities of peering, IP or optical (wavelengths on fibers) links. Our approach combines optimizations on single layers and *jointly* solves the following sub-problems: (1) assignment of end-user nodes to *hyper-giants'* ingress PoPs, (2) design of the IP topology (connectivity & capacity), (3) selection of optical paths for the IP links, and (4) routing of *hyper-giants'* and background demands in the IP topology. This modeling can leverage the flexibility of the three layers, e.g., demand can be routed via a dedicated IP link and optical path, or can use multi-hop IP routing to reduce resource fragmentation.

We start by describing the general joint optimization framework and then introduce specific optimization algorithms which come in different flavors, highlighting various aspects relevant in our evaluation.

A. Mathematical Model

Table I summarizes the sets and functions, Table II the constants, and Table III the decision variables.

Sets. The topology consists of two sets of nodes, IP nodes/routers (\mathcal{N}^{IP}) and optical nodes (\mathcal{N}^{Opt}), which are collocated and connected via optical transceivers. Multiple IP nodes can be collocated with one optical node, e.g., one for peering with *hyper-giants* and another for end-user connectivity. Some of the IP nodes provide connectivity for end-users (\mathcal{U}_h) or *hyper-giants* (\mathcal{P}_h) but not for both, i.e., $\mathcal{U}_h \cap \mathcal{P}_h = \emptyset \quad \forall h \in \mathcal{H}$, where \mathcal{H} is the set of *hyper-giants* and \mathcal{U}_h and \mathcal{P}_h are *hyper-giant*-specific sets. The set of pre-calculated equal-cost candidate paths between IP nodes e and f in the optical domain is denoted by $P_{e,f}$.

Constants. *Hyper-giant* demands are denoted as d_{uh} and peering capacity between ISP and *hyper-giants* as c_p^h in Gbps. The constant d_{ab}^{bg} indicates the amount of background traffic between IP nodes a, b . Each IP router has a maximum number

Table I: Sets and Functions

\mathcal{N}^{IP}	IP nodes (routers): At every PoP, there is at least one IP router.
\mathcal{N}^{Opt}	Optical nodes (ROADM): At every PoP, there is one optical node which is connected to the IP router at that PoP.
\mathcal{H}	Hyper-giants (HG): Entities that are responsible for the majority of the traffic.
\mathcal{U}_h	End-user demands of HG h .
$\mathcal{P}_h \subseteq \mathcal{N}^{\text{IP}}$	HGs' Peering locations: Routers where the ISP connects via PNIs to the HGs.
$P_{e,f}$	Equal-cost paths between $e, f \in \mathcal{N}^{\text{IP}}$ over the Optical Transport Network.
$i(u)$	$\bigcup_{h \in \mathcal{H}} \mathcal{U}_h \rightarrow \mathcal{N}^{\text{IP}}$ IP node where demand u destines to.
Q	A large number.

Table II: Constants

d_{uh}	HG Demand volume: Aggregated demand entering at router $i(u)$ towards the HGs in Gbps. Note that we allow multiple demands to the same HG from one $e \in \mathcal{N}^{\text{IP}}$. Different to others, e.g., [39], we do not differentiate between up- and down-link demands (request and reply).
c_p^h	Peering Capacity: Bandwidth of the PNI with h at $p \in \mathcal{P}_h$ in Gbps.
d_{ab}^{bg}	Background demand: Demand between IP routers $a, b \in \mathcal{N}^{\text{IP}}$ in Gbps.
t_e	Number of transceivers at IP router $e \in \mathcal{N}^{\text{IP}}$.
C	Capacity of a lightpath in Gbps.
w_{mn}	Fiber capacity: number of lightpaths that can be allocated on fiber between $m, n \in \mathcal{N}^{\text{Opt}}$.
u_{max}	$\in [0, 1]$ Maximum allowed IP link utilization.
l_p	Length of path $p \in P_{e,f}$.
ρ	Fraction of allowed reconfigurations.

Table III: Variables

$\hat{o}_{u,ph}$	= 1 if demand u to h flows over peering node $p \in \mathcal{P}_h$ to super sink h .
$o_{u,ef}^h$	= 1 if demand from u to h flows over IP link $e, f \in \mathcal{N}^{\text{IP}}$.
$o_{ab,ef}^{bg}$	= 1 if demand from a to b flows over IP link $e, f \in \mathcal{N}^{\text{IP}}$.
$y_{ef,p}$	IP trunk capacity: $y_{ef,p} \in \mathbb{N}$ indicates the number of lightpaths between $e, f \in \mathcal{N}^{\text{IP}}$ using path $p \in P_{e,f}$ and thereby the capacity ($\times C$) of the trunk.
r_{ef}	=1 if capacity of IP link changed in comparison to value from previous instance (y_{ef}^{t-1}).

of ports t_e which provide a capacity of C . The optical system restricts the number of lightpaths (wavelengths) on a fiber between m and $n \in \mathcal{N}^{\text{Opt}}$ to w_{mn} .

Variables. The variable $\hat{o}_{u,ph}$ indicates if demand u from *hyper-giant* h traverses peering node p , i.e., it represents the assignment of ingress PoPs to user demands. Another group of variables $o_{u,ef}^h$ indicates how the demands are routed in IP layer ($e, f \in \mathcal{N}^{\text{IP}}$), i.e., if the demand u from *hyper-giant* h flows over an IP link between e and f . Similarly, $o_{ab,ef}^{bg}$ indicates if background demand from a to b traverses IP link from e to f . We consider single path routing. Hence, all these groups of variables are binary valued.

The number of lightpaths that connect IP routers e and f while using path $p \in P_{e,f}$ is given by $y_{ef,p}$. As $y_{ef,p}$ already determines the routing in the optical domain, specific variables for optical routing are not required. Variable r_{ef} indicates if capacity of link $e, f \in \mathcal{N}^{\text{IP}}$ changed compared to the previous timestamp.

Constraints. Optimization is subject to several constraints

which address flow conservation, demand fulfillment and capacity limitations on three layers. Besides, limitations in the number of transceivers per IP node as well as bi-directionality of IP links and maximum link utilization are considered:

$$\sum_{p \in \mathcal{P}_h} \hat{o}_{u,ph} = 1 \quad \forall h \in \mathcal{H}, u \in \mathcal{U}_h \quad (1)$$

$$\sum_{f \in \mathcal{N}^{\text{IP}}: f \neq p} o_{u,pf}^h = \hat{o}_{u,ph} \quad \forall p \in \mathcal{P}_h, u \in \mathcal{U}_h, h \in \mathcal{H} \quad (2)$$

$$\sum_{f \in \mathcal{N}^{\text{IP}}: f \neq e, f \notin \mathcal{P}_h} o_{u,ef}^h - o_{u,fe}^h = 0 \quad \forall e \in \mathcal{N}^{\text{IP}}: e \neq u \wedge e \notin \mathcal{P}_h, u \in \mathcal{U}_h, h \in \mathcal{H} \quad (3)$$

$$\sum_{f \in \mathcal{N}^{\text{IP}}: f \neq u} o_{u,i(u)f}^h - o_{u,fi(u)}^h = -1 \quad \forall u \in \mathcal{U}_h, h \in \mathcal{H} \quad (4)$$

$$\sum_{u \in \mathcal{U}_h} \hat{o}_{u,ph} \cdot d_{uh} \leq c_p^h \quad \forall p \in \mathcal{P}_h, \forall h \in \mathcal{H} \quad (5)$$

Eqs. (1) through (4) describe the flow conservation for *hyper-giant* demands and thereby address sub-problems (1) and (4). Specifically, Eq. (1) sets the fractions of routed demand per *hyper-giant*-end-user pair to be equal to 1, i.e., all demands from *hyper-giants* (super-source) to end-users must be served. At all IP nodes where no end-users are connected, ingressing and egressing flow of a demand must be equal (Eq. (2) and Eq. (3)). At the destination node of a demand, it must sink (Eq. (4)). Eq. (5) limits capacity of peering nodes, i.e., the edges from the *hyper-giant*.

$$\sum_{f \in \mathcal{N}^{\text{IP}}: f \neq a} o_{ab,af}^{bg} = 1 \quad \forall a, b \in \mathcal{N}^{\text{IP}} \quad (6)$$

$$\sum_{f \in \mathcal{N}^{\text{IP}}: f \neq e} o_{ab,ef}^{bg} - o_{ab,fe}^{bg} = 0 \quad \forall a, b, e \in \mathcal{N}^{\text{IP}}, e \neq a, e \neq b \quad (7)$$

$$\sum_{e \in \mathcal{N}^{\text{IP}}: e \neq b} o_{ab,eb}^{bg} = -1 \quad \forall a, b \in \mathcal{N}^{\text{IP}} \quad (8)$$

Eqs. (6) through (8) are similar flow conservation constraints for background demands. They ensure that flows originate at the sources, terminate at the sinks and that in- and egressing volumes at intermediate nodes are the same.

$$\begin{aligned} & \sum_{h \in \mathcal{H}} \sum_{u \in \mathcal{U}_h} o_{u,ef}^h \cdot d_{uh} + \sum_{a,b \in \mathcal{N}^{\text{IP}}} o_{ab,ef}^{bg} \cdot d_{ab}^{bg} \\ & \leq C \cdot u_{\max} \cdot \sum_{p \in P(e,f)} y_{ef,p} \quad \forall e, f \in \mathcal{N}^{\text{IP}}: e \neq f \end{aligned} \quad (9)$$

$$y_{ef,p} = y_{fe,p} \quad \forall e, f \in \mathcal{N}^{\text{IP}}, p \in P_{e,f} \quad (10)$$

$$\sum_{f \in \mathcal{N}^{\text{IP}}} \sum_{p \in P(e,f)} y_{ef,p} \leq 2 \cdot t_e \quad \forall e \in \mathcal{N}^{\text{IP}} \quad (11)$$

Eq. (9) ensures that the demand flowing via IP link e, f does not exceed the given maximum IP link utilization and thereby,

addresses sub-problem (2). It considers only the edges adjacent to e . Eq. (10) and (11) account for bi-directionality of IP links, a lightpath is required for each direction, and bound the number of available transceivers per router (degree bound) respectively.

$$\sum_{e,f \in \mathcal{N}^{\text{IP}}} \sum_{p \in P_{e,f}: (m,n) \in p} y_{ef,p} \leq w_{mn} \quad \forall m, n \in \mathcal{N}^{\text{Opt}} \quad (12)$$

Eq. (12) limits the number of lightpaths that can be routed over a fiber and thereby addresses sub-problem (3). Note that the set of relevant paths, i.e., paths which use the optical edge m, n , can be pre-calculated.

$$\sum_{f \in \mathcal{N}^{\text{IP}}} \left[\frac{\sum_{h \in \mathcal{H}} \sum_{u \in \mathcal{U}_h: i(u)=f} d_{uh}}{C \cdot u_{\max}} \right] \leq \sum_{e,f \in \mathcal{N}^{\text{IP}}} \sum_{p \in P_{e,f}} y_{ef,p} \quad (13)$$

Eq. (13) provides a lower bound for the objective to reduce the run-time of the solver. In the ideal case, i.e., with minimum resource fragmentation, an IP node with end-user demands is connected via a single link to a peering node which is able to serve all the demands from this IP node. The necessary capacity of such a link is determined by summing the demands of all *hyper-giants* at that IP node and accounting for the maximum IP link utilization. The summed capacities of all links provide a lower bound for the total deployed capacity which is subject to fragmentation of link capacities and requires connectivity to several peerings of the *hyper-giants*.

$$\sum_{p \in P(e,f)} y_{ef,p} - y_{ef}^{t-1} \leq r_{ef} \cdot Q \quad \forall e, f \in \mathcal{N}^{\text{IP}} \quad (14)$$

$$y_{ef}^{t-1} - \sum_{p \in P(e,f)} y_{ef,p} \leq r_{ef} \cdot Q \quad \forall e, f \in \mathcal{N}^{\text{IP}} \quad (15)$$

$$\sum_{e,f \in \mathcal{N}^{\text{IP}}} r_{ef} \leq \rho \cdot |\mathcal{N}^{\text{IP}}|^2 \quad (16)$$

Eq. (14)-(16) limit the number of reconfigurations in terms of increased or decreased capacity per IP link. Eq. (14) pushes r_{ef} up if the capacity increases in comparison to the capacity of the previous timestamp y_{ef}^{t-1} . Eq. (15) addresses capacity decreases in a similar way. Eq. (16) limits the total reconfigurations over all IP links.

Objective. The main objective is to minimize the total deployed capacity which is given by the sum of capacities of all IP links:

$$\min \sum_{e,f \in \mathcal{N}^{\text{IP}}} \sum_{p \in P_{e,f}} y_{ef,p} \quad (17)$$

This objective function serves as a proxy for OPEX (e.g., power consumption) and CAPEX (e.g., transceivers to buy in long term). Additionally, it hints at the utilization of the optical topology that might in turn be used to operate λ -service during low traffic hours.

B. Greedy Ingress PoP Assignment

The MIP solves the sub-problems (1) - (4) simultaneously. In order to quantify the benefits of this joint optimization, we separate problem (1) and solve it with a greedy algorithm. The resulting PoP assignment is used as input for the MIP, which

solves sub-problems (2) - (4). We limit our evaluation to the separation of sub-problem (1) since it splits along the boundary between CDN and ISP. Moreover, this split first makes the demand more specific by fixing the source-destination-pairs for the *hyper-giant* demands and then optimizes the ISP's network. Other splits, e.g., optimizing the IP topology first, would counteract the demand-aware nature of our approach.

Algorithm 1 shows a greedy assignment procedure for a single *hyper-giant* based on the shortest path length between end-users and peering PoPs in the optical topology. The algorithm starts by sorting the end-users PoPs of this *hyper-giant* by their demand volumes in descending order (l. 1). For every end-user node u , it iterates over the peering nodes of this *hyper-giant* sorted by their distance to the end-user node in question (l. 4f). If the peering node p has enough remaining capacity and the number of transceivers suffices to accommodate the total assigned demand, u is assigned to p , i.e., $\hat{o}_{u,p} = 1$ (l. 6). Otherwise, the next peering node is evaluated. The algorithm fails if not all demands can be assigned. The procedure is repeated for all *hyper-giants*.

Algorithm 1 Greedy Ingress PoP Assignment for a single *hyper-giant* h .

Input: $c_p^h, d_{uh} \quad \forall u \in \mathcal{U}_h, p \in \mathcal{P}_h$

Output: $\hat{o}_{u,p}$

```

1: sort  $u \in \mathcal{U}_h$  by their demand volume  $d_{uh}$ 
2:  $allocDemand[p] \leftarrow 0 \quad \forall p \in \mathcal{P}_h$ 
3: for  $u \in \mathcal{U}_h$  do
4:   sort  $p \in \mathcal{P}_h$  by distance to  $u$ 
5:   for  $p \in \mathcal{P}_h$  do
6:     if  $allocDemand[p] + d_{uh} \leq c_p^h \wedge$ 
        $\left\lceil \frac{allocDemand[p] + d_{uh}}{C} \right\rceil \leq \frac{1}{2}t_p$  then
7:        $\hat{o}_{u,p} \leftarrow 1$ 
8:        $allocDemand[p] \leftarrow allocDemand[p] + d_{uh}$ 
9:     else
10:       $\hat{o}_{u,p} \leftarrow 0$ 
11:    end if
12:  end for
13: end for
```

C. Optimization Flavors

Combining different flexibilities for sub-problems (1)-(4) results in four different optimization flavors which we later compare in the evaluation. Note that for each flavor, we adapt the MIP accordingly, i.e., fix variables:

Baseline uses ingress PoP assignment, IP connectivity and routing from the real data, without performing any additional optimization; it determines the capacity of the IP links to serve all demands. Note that the topology might include some requirements for particular links due to business policies.

ISP-only uses the ingress PoP assignment extracted from the real data and optimizes the IP topology and routing with fixed PoP assignment. This ISP-only approach reflects a scenario where the CDNs are not following the cooperation.

2-step optimizes the PoP assignments using Algorithm 1 and uses the optimal approach with the PoP assignment for IP

topology and the routing. Thus, CDNs are cooperative but optimization is done separately in two steps.

Joint jointly optimizes all three layers, PoP assignment, IP topology and routing, using the exact approach from §III-A.

IV. EVALUATION

This section empirically answers our fundamental question: can we leverage reconfigurable topologies to improve ISP traffic steering and resource management? To do so, we first outline the attributes of the optical infrastructure, the traffic and the optimization approaches (§IV-A). We then assess to what extent we benefit from reconfigurations – if at all – and what frequencies of reconfigurations are needed (§IV-B). We complement this analysis by evaluating the benefit of jointly optimizing along all available dimensions (§IV-C) and a dissection of the costs in terms of observed reconfigurations concerning the topology's recurrence and stability (§IV-D). We conclude by considering how our results generalize to specific scenarios. To this end, we investigate our approach in the context of the traffic workload during the COVID-19 pandemic, a scenario with link failures, and a scenario with randomized demands (§IV-E and §IV-F).

A. Settings

The evaluation relies on real topology and traffic data collected from a large ISP with > 15 million fixed line and > 30 million mobile users. We limit data to the network in the ISP's home country which consists of more than 10 Points-of-Presence (PoPs) and serves more than 50 PB of daily traffic.

Optical infrastructure. The optical network (ON) consists of < 20 nodes and < 30 edges. Each edge has an optical capacity with $w_{mn} = W = 100$ lightpaths (wavelengths). Connectivity between access routers of the end-users and core routers is fixed, and optical reconfigurability is limited to the core layer. Hence, we aggregate the access routers into the corresponding core routers. Thereby, we reduce the original set of IP routers (> 1k) to the core routers and the largest peering routers resulting in < 30 IP routers in total. While this reduction is rooted in the capabilities of the network, additionally, it also makes the optimization more tractable. At some PoPs multiple IP nodes are connected to the optical node. Every router can support $t_e = 100$ transceivers. A single lightpath has a capacity of $C = 100$ Gbps. Moreover, the maximum IP link utilization is limited to $u_{max} = 50\%$ to accommodate traffic in cases of failures. The *hyper-giants'* peering locations and capacities are also extracted from the ISP's data.

Traffic (Demands). To create traffic demands, we first collect real traffic data from the FlowDirector system [16]. FlowDirector gathers > 45 billion NetFlow records per day at the border routers of the ISP. Then, we differentiate two types of demand samples: HT and BT.

For HT, a demand sample contains only flows belonging to the top-10 *hyper-giants* in the ISP network. It makes > 75% of the total traffic in the ISP's backbone network. In order to create the input for our model, the flows are aggregated. For HT, the aggregation is done based on the *hyper-giant*, the origin peering router, and destination core router of the

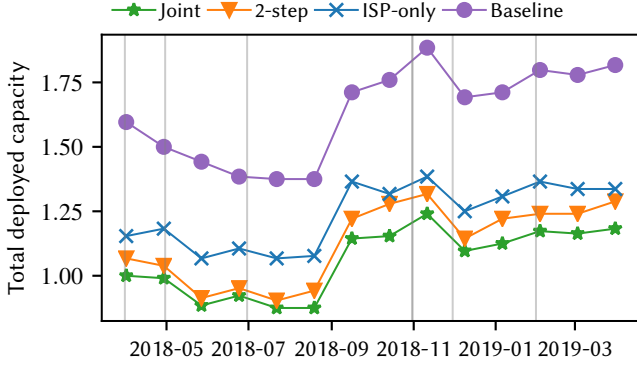


Figure 2: Total deployed capacity with reconfigurations on a monthly basis for HT-only. Gray lines indicate events where peering for at least one of the *hyper-giants* changed.

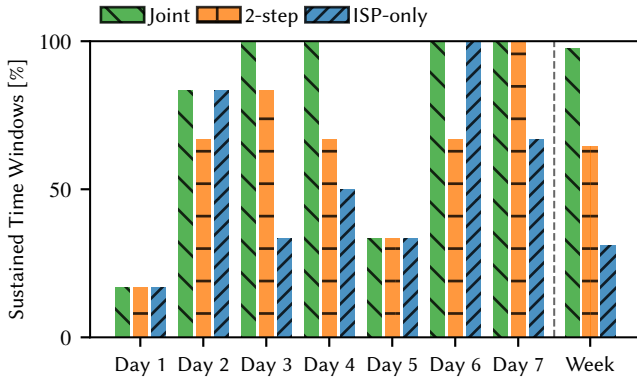


Figure 3: Re-optimizing once per day/week with full knowledge about peak hour. Bars indicate fraction of time windows where the day's/week's peak hour solution is feasible.

flows. For background traffic (BT), we select the traffic that does not belong to any of the top-10 *hyper-giants*. Then, we aggregate flows based on their origin and destination core routers. Further, as described later (§IV-B), the demands are split and aggregated according to the time windows, e.g., two hours or one day. Therefore, the average rate per demand over one hour is calculated. Finally, the maximum value of the time windows is used as demand value in the optimization. We use either HT-only or AllTraffic, the combination of HT and BT, as input to the optimization.

Optimization approaches. The main performance objective is the deployed capacity in the IP topology which serves as a proxy for OPEX (e.g., power consumption) and CAPEX (e.g., transceivers to buy). The evaluation examines the optimization approaches from §III-C and a theoretical baseline (*Theo. min.*). This baseline calculates the deployed capacity for the case where all end-user nodes are directly connected to only one peering node (CDN PoP), building a star topology. Such a star topology requires that there are PoPs with sufficient capacity where all *hyper-giants* peer with the ISP. An assumption that does not generally hold as the set of peering locations usually differs among the *hyper-giants* [16]. Moreover, *Theo. min.* ignores capacity limits of the fibers and number of transceivers per router.

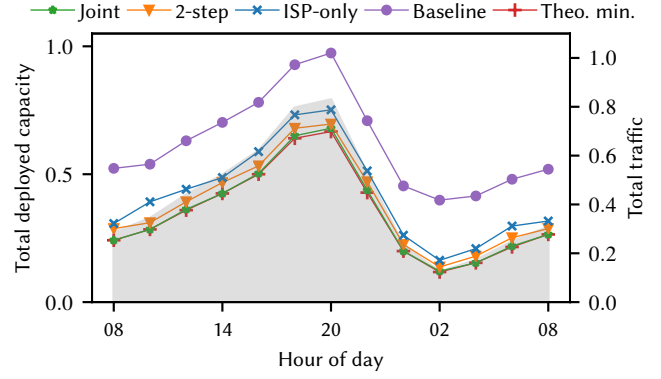


Figure 4: Re-optimization on bi-hour-scale. Shaded area indicates total traffic volume normalized to maximum value from AllTraffic.

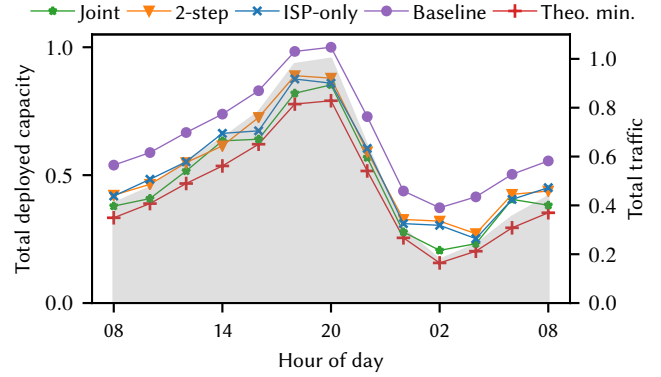


Figure 5: Re-optimization on bi-hour-scale for AllTraffic case. Shaded area indicates total traffic volume normalized to maximum.

If not stated otherwise, each optimization determines all components of the solution (PoP assignment, IP topology). We use the Python API of IBM ILOG CPLEX 12.9 [40] to implement the MIPs. The solver time limit is set to 1 hour.

B. Reconfiguration Intervals

1) *Monthly-based reconfigurations:* According to a previous study, the collaboration between CDNs and ISPs is carried out on a daily to monthly basis [16]. Hence, we believe that investigating reconfigurations on a monthly basis is a good starting point. In this case, re-optimization is performed once per month on maximum values of the weekly peak-hour demands of that month. We investigate the traffic from April 2018 to April 2019 and focus on HT-only due to space restrictions.

Fig. 2 shows the total capacity normalized to the value of Joint at March 2018. The gray vertical lines indicate dates when the peering of at least one of the *hyper-giants* changed, i.e., peering capacity increased/de-creased or its location changed. Here, we might see whether re-optimization benefits both *hyper-giants* and ISPs.

Reconfiguration leads to capacity savings. For all algorithms, the trend of the total capacity increases over time. Baseline performs significantly worse ($\approx 35\%$) on average than all other algorithms. The benefit of ISP-only

comes from changing IP link capacities more aggressively and adapting the IP routing in contrast to Baseline. Going a step further and also including CDN user mapping into the optimization, Joint saves up to 15% in comparison to ISP-only. Doing a greedy mapping here (2-step) still saves 5–10% in comparison to ISP-only. Overall, there is no clear trend of benefiting from peering changes for all approaches. To summarize, introducing monthly reconfigurations leads to substantial savings in infrastructure upgrades. However, the question remains whether such optimization can really sustain all traffic changes within a month: for instance, events can lead to severe traffic changes at specific locations, which might be obscured by the aggregation with the global maximum. Hence, we look into this by considering the traffic demands of a week.

2) *Weekly-based and daily-based reconfigurations*: We divide a week into four hour long time windows and use the maximum demand hour of each window as input. Here, we again assume to know the peak hour per day/week in advance and run three optimizations: Joint, 2-step, ISP-only. We want to answer the question whether the peak hour's topology sustains all traffic patterns of this week while keeping the maximum IP link utilization constraint. Fig. 3 visualizes the fraction of time windows sustained by the peak traffic hour solutions.

Coarse reconfiguration intervals do not sustain traffic. The result, however, is negative: for all three week-peak-hour-based solutions, the fractions are < 100%, i.e., there is at least one point in time when the solution cannot satisfy the end-user demands. Similarly, the results for daily re-optimizations are shown. The result is again negative. There are several days of the week for which daily re-optimization is not sufficient. This demonstrates the need for reconfigurations on smaller time-scales.

3) *Hourly-based reconfigurations*: Fig. 4 shows the total deployed capacity throughout a day. Optimization is performed periodically every two hours based on the traffic maximum of the considered time-period. The gray area contrasts the total traffic volume. The lines and markers indicate the values of the best integer solutions found within the time-limit of the solver. Values are normalized to the maximum values from the AllTraffic case, i.e., maximum value of Baseline for capacity values and maximum value of traffic volume for total traffic.

Intra-day reconfigurations save up 44% of capacity over time. As before, all three algorithms save consistently more than 30% of capacity in comparison to Baseline. Moreover, they have a consistent order throughout the day: Joint performs best with an average gap of 3% to *Theo. min.* followed by 2-step and ISP-only with gaps of 11% and 25%. Thus, jointly optimizing ingress PoPs routing and topology can save non-negligible fractions of capacity.

In addition to the savings compared to the Baseline, reconfiguration on a bi-hourly-bases also opens the possibility to exploit the diurnal demand patterns: The ratio between peak hour (around 20:00) and the deepest valley (around 02:00) is ≈ 7 for both traffic cases. This variance is also reflected in the deployed capacity for all algorithms. Re-optimizing the

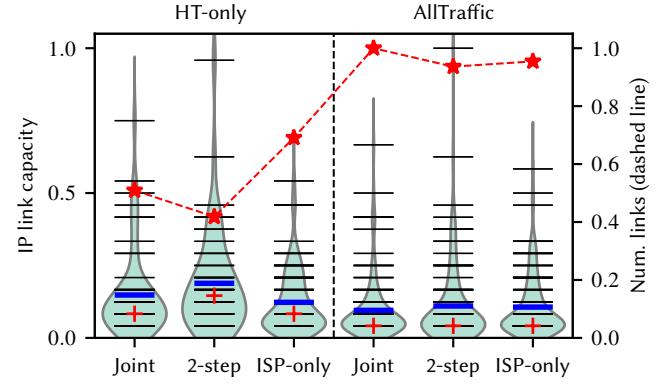


Figure 6: Distribution of normalized IP link capacities and normalized number of links (dashed line).

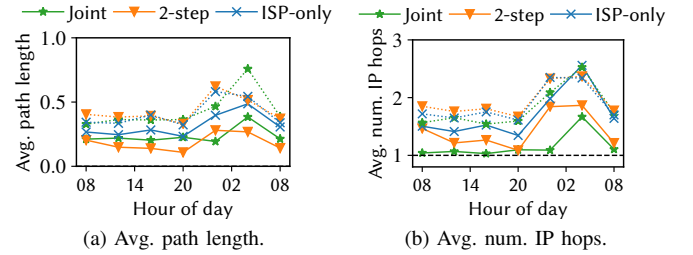


Figure 7: Comparison of average path length (normalized distance) in arbitrary distance unit and average number of hops for HT (solid) and BT (dashed).

network saves 44% of the summed capacity over the day. This directly translates to reduced OPEX, e.g. by energy savings, or freed fiber capacity that can be re-used for other services. Hence, there is a need for reconfiguration on an hourly-basis in order to efficiently use the capacity.

C. Need for Joint Optimization and Reconfigurations

In this section, we answer the question how joint re-optimization of the IP-topology and PoP assignment can serve ISPs and *hyper-giants*. Further, we provide insights into the efforts in terms of reconfigurations. We consider both HT-only and AllTraffic, i.e., we always highlight the differences between both traffic types.

1) *The benefits of the ISP and the hyper-giants*: Whereas the previous section showed that we can generally save capacity when considering *hyper-giant* traffic, this section takes a deeper look into the impact also of the BT. BT accounts for $\approx 25\%$ of traffic volume (Fig. 4 vs. Fig. 5). While Baseline does not experience a significant rise in capacity, the other approaches account for this increase in traffic. The differences diminish: Joint still achieves the best performance, but 2-step and ISP-only come closer. Nevertheless, re-optimization still exhibits 15% (in contrast to 30% for HT-only) better performance than Baseline—reconfigurations can still lead to significant savings for ISPs.

We now deepen the analysis of the differences among the optimization flavors. For this, we look at the attributes of the solutions, i.e., their link capacities, the average path lengths, and the average number of IP hops.

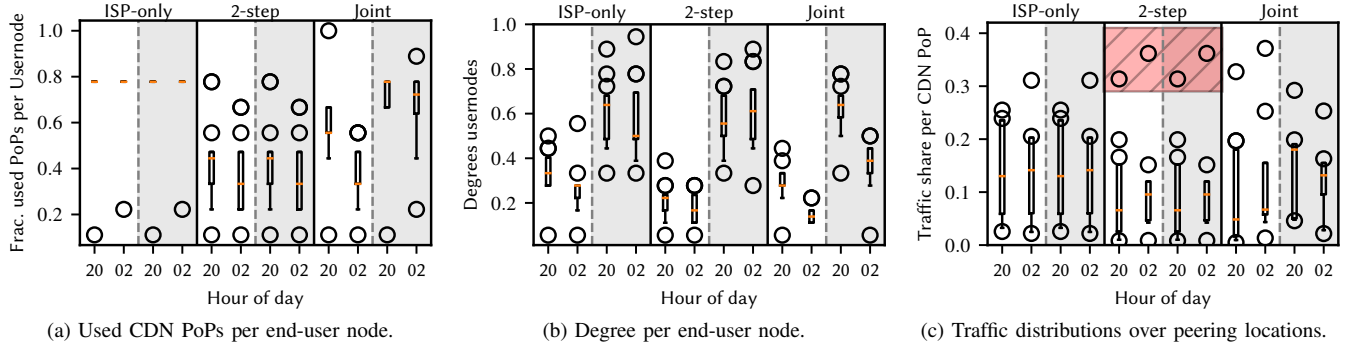


Figure 8: Used CDN PoPs by each approach (a), degree of IP nodes by each approach (b) and traffic distribution over peering locations (c). Optimization scenarios: HT-only (white area) and AllTraffic (gray area).

Joint optimization creates smaller topologies. Fig. 6 compares the number of IP links and the distribution of their capacities for all algorithms with HT-only and AllTraffic; here, we take the sample at 18:00. For HT-only, significant differences between the algorithms are visible: 2-step deploys the smallest number of links followed by Joint and ISP-only. But it has significantly larger link sizes (the maximum capacity is almost twice as that of ISP-only). Moreover, ISP-only has the largest number of very small links as indicated by the shape of the violin. This is inefficient from the capacity point of view as it results in large excess capacities but increases connectivity. Joint is between these two.

For AllTraffic, the differences in numbers reduce, as for all algorithms the number of links significantly increases. However, the medians of the capacity distributions (red cross) drop for all algorithms which means that the added links have mainly small capacities to add the necessary connectivity for BT. The reduced number of links for Joint results in shorter paths for HT as described in the following.

Joint optimization decreases mean path length. One primary goal of *hyper-giants* is an efficient delivery of their traffic towards end users. This does not only manifest in high throughput, but also in small latency values. The average path length measured either in IP hops or real distance (arbitrary unit) are two ways to assess this aspect: for instance, the shorter the paths, the lower the expected latency.

Fig. 7a contrasts the mean path lengths for the AllTraffic scenario. When comparing the average path length of the HT and the BT, we observe that HT is consistently routed via shorter paths (average lengths of 0.3) than BT (average lengths of 0.45). The algorithms themselves do not clearly impact the average path lengths: for instance, 2-step is sometimes better and sometimes worse than Joint.

In contrast, when looking at the average number of IP hops (Fig. 7b), Joint shows a superior solution: it always has the least amount of average IP hops for the HT which reduces the total required capacity in the network (bandwidth tax). Here, ISP-only suffers from the PoP assignment strategy; their average path lengths are $\approx 50\%$ higher — a clear benefit of the joint optimization. As elaborated later, ISP-only has a very diverse PoP assignment which results in longer path to reduce the deployed capacity. On the other hand, Joint is

able to determine such an PoP assignment that most traffic is routed via direct links. For BT, the differences among the algorithms are less strong.

In order to understand the solutions (i.e., how the “designed” topologies look like), we compare the approaches with respect to the used CDN PoPs and the IP node degrees. For the used PoPs, we look at each end user node and determine the number of PoPs this end user node addresses. For the degree of the IP nodes, we determine the number of links to other IP nodes.

Aggregated end-user mappings reduce IP node degrees by 10%. Fig. 8a shows boxplots of the number of peering locations (CDN PoPs) that are used by the end-user nodes for the peak (20:00) and valley (02:00) times of the traffic. ISP-only uses the mappings from the collected data and shows that most end-user PoPs have *hyper-giant* demands routed via almost all CDN PoPs, i.e., demands are less aggregated. For 2-step and Joint, the fractions of used PoPs are smaller and vary throughout the day. The rational behind this is that both algorithms try to group the demands and route them via the same PoP and path in the topology to reduce fragmentation of IP capacity. As the size of the demands increases towards the peak hour and peering capacities are limited, mappings have to be adapted, become more distributed, and more PoPs are used during the peak hour. Fig. 8b illustrates that this behavior leads to reduced nodal degrees of the end-user nodes. Recalling the ideal situation, where all traffic of the end-user node is routed via a single link, this is the desired behavior and the potential that is exploited by the joint optimization. For AllTraffic, the differences in node degree are smaller as also observed for the total deployed capacity (cf. Fig. 5).

2) *Are there any drawbacks for hyper-giants?* The optimization in this work focuses on cost reductions for ISPs. Although this leads to shorter distances between *hyper-giants* and their end users, this might affect the utilization of the resource provisioning of *hyper-giants*. Hence, we look at the traffic share of the *hyper-giants*’ PoPs, a potential indicator for the PoP utilization. We believe that *hyper-giants* rather prefer evenly and constantly utilized peerings (better load distribution and less changes over the day).

Utilization of peerings is unbalanced. Fig. 8c shows the allocated traffic share over the peering locations of the *hyper-*

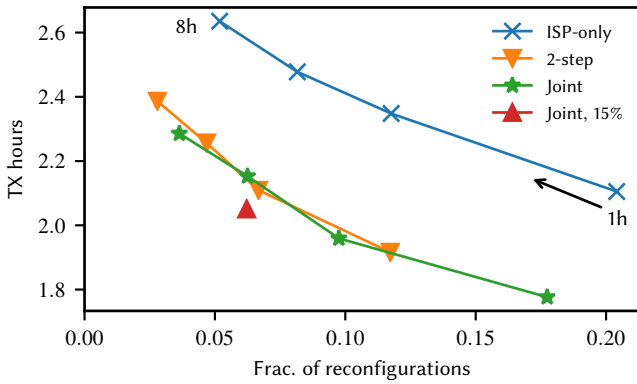


Figure 9: Pareto plot of reconfigurations vs. transceiver (TX)-hours (both normalized); optimization traffic is HT-only.

giants. For ISP-only, load distribution values are similar for both HT-only and AllTraffic. This is an expected behavior, as the *hyper-giant* demands do not change for both optimization scenarios – but it proves the operation of the optimization approaches. Similar observations hold for 2-step. However, the overall traffic share is lower for the majority of PoPs. In order to sustain all demands, 2-step uses a PoP with a high traffic share, as shown by the outlier above a value of 0.3 (hatched box). Baseline is left out as it has the same results as ISP-only.

In contrast to ISP-only and 2-step, Joint shows changes in distribution and variability during the day. Values range from 0.1 to 0.5. As a consequence, *hyper-giants* might have to adapt their peering capacities, and potentially also their routing towards peering points, over the day more frequently – a potential trade-off for *hyper-giants* between operational efforts and lower latency values.

D. Reconfigurations: Analyzing Topological Impact

So far, our conclusion is positive: joint reconfigurations help to reduce needed capacity for ISPs and shorten the path length from end-users to *hyper-giants*. On the other side, operators have been (and still are) reluctant to use frequent, short-term reconfigurations in their network. Accordingly, in this section, we study (1) the impact of the total number of reconfigurations needed and (2) which type of reconfigurations drive the capacity gain. Based on such insights, operators can better trade off the costs in terms of reconfigurations and capacity gains according to their preferences.

1) *Trade-off between reconfigurations and capacity deployment*: As a first simple solution, we study the impact of longer reconfiguration periods on the total capacity; the different periods are 1, 2, 4 and 8 hours. As in §IV-B, we optimize for the peak demands of the time windows.

Large optimization periods reduce reconfigurations. Fig. 9 shows a Pareto plot between the normalized number of reconfigurations and the transceiver (TX) hours. TX hours represent the integral of the deployed capacity over time, normalized by *Theo. min*. As expected, longer periods reduce the amount of reconfigurations but they increase the spent TX hours. Among the algorithms, ISP-only performs worst in both dimensions. 2-step results in less reconfigurations

compared to Joint at the cost of more TX hours which indicates one way to trade-off both objectives. The savings in reconfigurations reduce with increasing periods.

Considering Joint, we note significant savings ($\approx 45\%$) in reconfigurations when re-optimizing every 2 hours instead of 1 hour, while the TX hours increase only by $\approx 10\%$. For 8 hours, Joint has the least number of reconfigurations but also a significantly higher capacity (22%). This describes the landscape where operators can trade-off the TX hours savings against reconfigurations.

Consequently, we add a *limiting reconfiguration* constraint (cf. Eq. (16)) that makes the Joint approach a credible alternative for 2-step. The figure shows that optimizing every 2 hours and limiting the number of reconfigurations to 15 % can provide better results in terms of reconfigurations and TX hours than 2-step. Our proposed model does not only demonstrate the benefit of joint optimization of IP topology, routing, and PoP assignment, but also provides flexibility to adjust results based on, e.g., provider policies.

2) *Impact of different reconfiguration types*: Fig. 10 presents a breakdown of the reconfigurations into their types: IP links added, IP links removed, IP links with capacity increased, and IP links with capacity decreased.

Adjacency changes impact only small share of the traffic.

Fig. 10a shows the normalized number of reconfigurations for all types. While $\approx 50\%$ of the reconfigurations for Joint and ISP-only are adjacency changes, the share is smaller for 2-step. Also Joint in combination with the reconfiguration limit mainly impacts link sizes. Here, the algorithm avoids changes to the adjacency matrix of the IP topology and adaptations are mostly facilitated by changes in capacities. Note, however, that capacity savings cannot be achieved without any link adaption. Assessing the size of the modified links, Fig. 10b illustrates the link capacities before the change (and after the change for additions). While there is no clear pattern among the algorithms for scaling links up or down, additions and removals are restricted to very small links for ISP-only, 2-step, and Joint with limited reconfigurations. This suggests, that only minor fractions of the traffic are affected by these changes and hints towards a stable topology structure.

End-user re-mappings foster capacity savings. Fig. 10c visualizes the changes in the PoP assignment and in the IP routing when the CDN PoP for the demand stays the same. The observed behavior in §IV-C1 results in high numbers of end-user mapping changes and according routing changes for the demands for Joint without (∞) and with limit (15%) on the reconfigurations which represent another factor to trade-off for the reduced capacity. 2-step and ISP-only show more static mappings but still high numbers of changes in the routing of the demands. This demonstrates how reconfigurations of the CDN user mapping foster the capacity savings.

3) *Stability of topology*: Motivated by the observation that mainly small capacity links are removed and added to the topology (cf. Fig. 10b), we ask to what extent there is a persistent structure in the topologies.

Large fraction of traffic is routed via persistent links. Fig. 11 shows fractions of links (Fig. 11a) and carried traffic

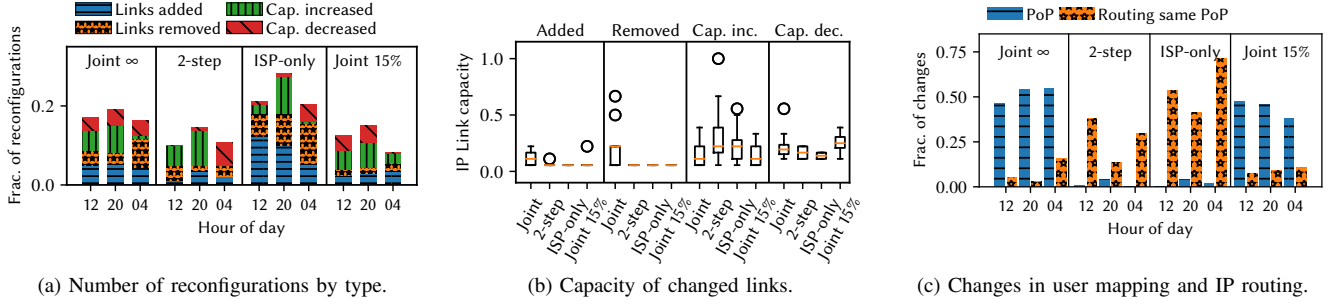


Figure 10: Details on reconfigurations by type, algorithm and affected entities for HT-only. *Joint* which deploys least capacity in the network, requires more and more evolved types of reconfigurations.

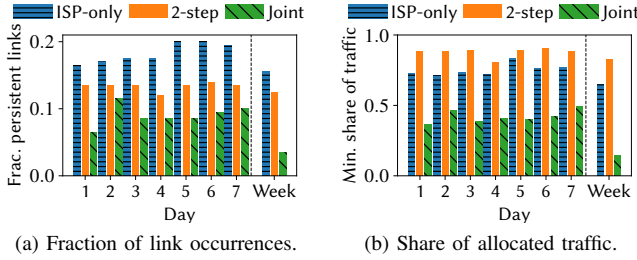


Figure 11: Stability of topology. (a) compares occurrence of unique links over one day and one week. (b) shows the share of traffic over persistent links.

(Fig. 11b) for persistent parts of the topology, i.e., for links that exist throughout multiple optimization periods. The figures evaluate two time windows for persistence: per day and per week. For *2-step* and *ISP-only*, the fraction of adjacencies that exists over the whole day is consistently $> 10\%$ and carries between 70 – 90 % of traffic. For *Joint*, the persistence is less strong on a day-by-day basis (5–10 % of the links) which aligns with the observations of reconfigurations of larger links. Also the fraction of carried traffic over this stable topologies is smaller but with 40 % still significant. Considering the span of a week, for *Joint*, less than 5 % of the links exist all the time. For *2-step* and *ISP-only*, the fractions are still $> 10\%$ and the links carry $> 60\%$ of the traffic. These observations underline that most reconfigurations affect only small fractions of the traffic and that significant parts of the traffic are routed via stable parts of the topology that are only scaled up and down. Moreover, setting a stable topology on a daily basis and temporarily adding small links upon short term spikes, allows to maintain higher operational confidence while gaining from reconfigurability.

E. Special Event COVID-19 and Failures

We evaluate the benefits of our approach in situations with increased and changed traffic patterns. As a case study, we use data collected during the recently active COVID-19 pandemic. The samples are aggregated from 4 hour windows on three different days. *BEFORE* is a day in the week just before the major lock-down in the country of the ISP, *AFTER 1* and *AFTER 2* are the same weekday in the two subsequent weeks.

Fig. 12a shows the deployed capacity during peak hour for the case of unlimited reconfigurations. For all algorithms,

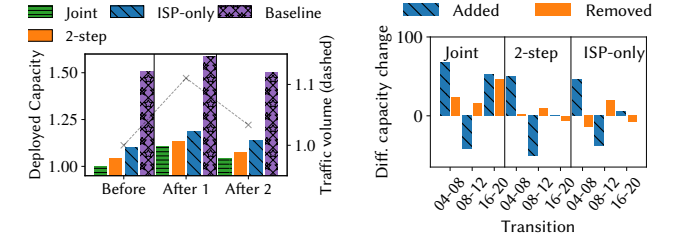


Figure 12: Impact of traffic during COVID-19 on deployed capacity and reconfiguration behavior.

the deployed capacity increases for *AFTER 1* and *AFTER 2* in comparison to *BEFORE* by $\approx 13\%$ and $\approx 4\%$ respectively. This is in line with the increase in traffic volume observed [26], [41], [42]. The decrease of deployed capacity and traffic volume for *AFTER 2* aligns with the decrease of video streaming quality by two *hyper-giants* [43]. The already observed behaviors of the algorithms from sections before are not affected significantly by the traffic increase with *Joint* being $\approx 10\%$ better than *ISP-only* and $\approx 2\%$ better than *2-step*.

Events impact reconfiguration behavior. Fig. 12b contrasts the changes in capacity between *AFTER 1* and *BEFORE* for several transitions throughout the day. It shows the differences in summed capacity of all links added or with increased capacity and all links removed or with decreased capacity. For all algorithms, *AFTER 1* has higher capacity increase from 04:00 to 08:00 compared to *BEFORE* (difference > 0), while the capacity added in the transition from 08:00 to 12:00 is lower. Only for *Joint*, we also observe higher differences in the amount of removed capacity which indicates more changes in connectivity. Thereby, *Joint* can use the ISP's optical infrastructure and CDN peerings more efficiently.

Joint IP routing and user mapping restoration increases robustness upon IP link failures. Although our approach does not directly optimize the network for resilience, we evaluate the ability to restore from single failures out of the 25 largest IP links of the topology. Such a scenario reflects failures of transceivers. To evaluate the worst case scenario, we assume that 100% of the link capacity becomes unavailable. To start with, we focus on two restoration possibilities: IP routing and CDN PoP assignment and keep the IP topology

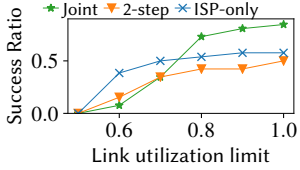


Figure 13: Restoration success ratio for single IP link failures against allowed link utilization.

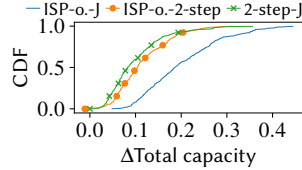


Figure 14: CDF of relative difference in deployed capacity between optimization flavors for randomized demands.

fixed. Note, that we do not optimize for resilience, hence the following analysis provides only an outlook.

Fig. 13 compares the ratios of successful restorations of the three algorithms against the allowed link utilization. While Joint shows increasing restoration success with more relaxed link utilization and eventually can restore 80 %, 2-step and ISP-only saturate quickly around only 50 %. The distribution of the IP link capacities (Fig. 6) reveals that ISP-only deploys more links but the single links' capacity is smaller. This in turn leaves only little headroom to allocate traffic in case of re-routing. On the other hand, 2-step deploys few large links, which carry high traffic volumes. Failures of those IP links require high excess capacity for restoration, making it less flexible. We acknowledge that not all failures can be restored but leave consideration for future work.

F. Randomized Demand Patterns

Finally, we assess the generalization of our results to randomized demand patterns. As described in §IV-A, the data has a specific structure, and hence, it is challenging to generate randomized demands that are feasible. For instance, given the peering capacities of the *hyper-giants*, uniform sampling of demands can lead to situations where demands exceed the available capacity of the peerings, or where a feasible routing cannot be found. Therefore, to avoid infeasible problem instances, we scramble the existing input data along the spatial dimension. That is, we shuffle the assignment of IP end-user nodes and peering nodes to the optical nodes. Thereby, the relations between peering capacities and demand sizes are preserved.

Fig. 14 shows the cumulative distribution functions of the relative differences between the algorithms. It uses the input data from Fig. 4 and randomizes it with 30 seeds – 360 samples in total. Joint (“ISP-o.-J”) and 2-step (“ISP-o.-2-step”) both perform better than ISP-only. For Joint, all differences are > 0 with an average of 0.2, i.e., a 20% improvement. 2-step gains 11% on average compared to ISP-only on the randomized data. Moreover, the difference between 2-step and Joint (“2-step-J”) is ≥ 0 for all samples. In 75% of the instances, Joint can save more than 5% of deployed capacity compared to 2-step, and 40% of the instances show savings $\geq 10\%$. Overall, this demonstrates that the gains of joint re-optimization generalize also to less structured demand patterns.

V. DISCUSSION: REAPING THE POTENTIAL BENEFITS

Our evaluations based on data from a large ISP revealed a significant potential of a joint optimization and adaptation of and for *hyper-giants*' traffic – for the ISP in terms of reduced costs or increased resource efficiency, and for the *hyper-giants* in terms of reduced latency. In the following, we discuss limitations and avenues on how these benefits may actually be reaped.

A. Generalization and Scalability

Our evaluations consider only one ISP topology, a limitation of our work that stems from the lack of publicly available data for other topologies (cf. §IV-F). Nevertheless, we expect the results to generalize to other ISP networks as well since those networks share important properties, at least to those of comparable size. First, similar traffic patterns like the diurnal pattern and dominance of *hyper-giants*' traffic in ISP networks have also been observed for other ISPs [44]. Traffic demands increase and patterns might change, as also observed during the COVID-19 pandemic [42]. Our sensitivity analysis with synthesized demands confirms our observations and provides us further confidence in the robustness of the results. A second aspect of the generalization is whether the optimization is tractable on other ISPs. Our modeling results in an NP-complete problem which might lead to run-time problems on larger problem instances. Limiting the solver run-time to 1 hour has provided results close to the optimum in our case but might not be sufficient on other topologies or even faster reconfiguration cycles. Hence, designing more tailored algorithms might be necessary for real deployments. We leave this for future work.

B. Business-related Trends

Emerging cooperation. Many inefficiencies related to how CDN traffic is delivered through ISP networks today, are due to the lack of information. While ISPs have detailed knowledge of their network topology (the two lower layers in Fig. 1), their knowledge of the demand of content and distribution flexibilities is usually very limited (upper layer in Fig. 1); and vice versa for the *hyper-giant* [21], [45], [46]. Several recent studies presented promising approaches to improve information sharing and cooperation [16], [21], [47]–[50]. Moreover, it has been shown that major players adopt such approaches for information exchange [16], which we argue can and should be extended to account also for topological flexibilities.

Hyper-giants acting as ISPs and vice versa. In addition to emerging collaborations, we can already note that some *hyper-giants* become ISPs and provide connectivity for end users. Either on traditional fixed line access [51] or via less established interconnects such as satellite links [52]. Similarly, ISPs started to provide content services, e.g., [53], [54]. There is an increasing number of entities unifying ISP and *hyper-giant*. Hence, our approach of joint optimization is becoming more relevant.

C. Technological Trends

From a technical perspective, our approach largely benefits from the increasing adoption of softwarization and centralized control approaches in ISP networks, e.g., [11]. In particular, there are two major aspects to realize the envisioned system.

Centralized data collection and operation. The joint optimization relies on detailed knowledge of the current state of topology and routing, and particularly on information about the future demands. Systems that implement such data collection in a highly scalable way are already in operation at the ISP, e.g., [16], using widely deployed protocols such as BGP or IS-IS for routing data and NetFlow or IPFIX for collection of demand data. Prediction of future demands based on the collected ones is already possible with deployed systems [10] and fostered by omnipresent advances in machine learning. Also data from the CDNs, e.g., server loads or content availability, is already being collected by CDNs to implement their user-mapping systems [33]. Besides the centralized database, also the control and decision making has to operate on a global level to leverage the joint optimization. Large ISPs and *hyper-giants* have shown that centralized control entities are feasible already today, cf. [2], [3], [11]. Depending on the particular situation, i.e., CDN-ISP cooperation or implemented by one party, detailed integration of all parts into one system is necessary. However, further work to integrate control over the three layers is still necessary.

Deploying configurations. Finally, configurations have to be communicated to the network equipment in case of IP topology and routing or to the mapping systems for content requests. Run-time reconfigurations of network devices have strongly been fostered with the continuing trend of softwarization and programmability of networks. Most devices offer programmable interfaces to change configurations and the state of the network. For instance, in the optical domain, NetConf and TL1 are prominent examples for such interfaces. But often the specifics are vendor dependent potentially posing a challenge for system integration. Note that while such interfaces offer easy triggering of changes, actual realization by the device might still be limited by other technological aspects. For the deployment of user mappings, approaches depend on the involved entities. For CDN-ISP collaboration, communication between the parties should happen in a reliable and automated way. ALTO [37] has already been adopted by some players to achieve this [16]. Other approaches might include custom APIs between mapping system and centralized control, e.g., in the case of converged roles of CDN and ISP. Reconfigurations can lead to interruptions or additional overhead, e.g., if states have to be synchronized over the network or if IGP has to re-converge. A specific detail here is the scheduling of reconfigurations to reduce service interruptions and to guarantee correctness of the networking state during reconfigurations, e.g., using a *make-before-break* strategy [55]. The specific costs depend on the actual system design. A deeper analysis particularly for reconfiguring on all three layers is subject to future work.

VI. RELATED WORK

To the best of our knowledge, this paper provides the first study of how reconfigurable topologies can be used to improve CDN traffic routing in ISPs. Our paper builds upon several important existing results in the area of optical networks, CDN-ISP collaboration, and network resource optimization, which we will discuss in turn in the following.

Capacity planning and routing in optical networks. Capacity planning is a classic problem in optical networks and there already exists a large body of literature, also accounting for multi-layer aspects. For a survey, we refer to [13], [56], [57]. Much existing related work on optical network optimization revolves around the impact of optical network reconfiguration on the routing stability [58], [59], the reconfiguration and adaption of virtual topologies that overlay optical networks to changing traffic demands [60], issues related to incremental deployment [14], [61], as well as regenerator placement problems (i.e., computing ROADM locations). All these works consider the demand given by source-destination pairs and do not shape it via optimizing end-user mappings like our proposed approach does. There is also interesting work on content-oriented and application-aware optical network optimization [39], [62], and multi-layer resource allocation [15], [63], [64], e.g., in the context of the Facebook network [65]. A case study related to CDNs is conducted in [66], however, without considering reconfigurations. In particular, these works consider optimizations on the demand level but optimize only two layers neglecting flexibility of IP grooming [39], [66], and they do not adapt over time [39], [62], [65]. Adaptive operation is provided by [15], [63] along with integration of application requirements. However, the demand is given again by fixed source-destination pairs. While there exist several efforts to render operations of optical wide-area networks more adaptive [9], [67], we are not aware of any related work on the optimization of ISP networks along all the three dimensions arising in the context of CDNs: topology, routing, and end-user mapping. In particular, the few existing two layer solutions which account for multiple mapping locations, such as [39], do not support IP grooming. We also note that while our focus in this paper is on ISP topologies, the opportunities introduced by reconfigurable optical networks has recently also received much attention in datacenters, e.g., [68]–[71]. However, the technologies and constraints in datacenters are fairly different from the ones in ISPs (e.g., in terms of routing [72], availability of wireless channels [68], and concerning the workloads [73]), and existing results are not easily transferable.

CDN-ISP collaboration. There is interesting work on how collaborations between CDNs and ISPs can be improved. This work, however, mainly focuses on traffic engineering (TE), neglecting the potential of topology modification. For related work on joint content placement and TE in this context see [47]–[49], [74]–[76]. All this work is limited to user mapping and TE but ignores adaptive optimization of the ISP's topology. A recent case study considers the joint content distribution and TE of adaptive videos in telco-CDNs [77], assuming a caching system where CDNs can deploy media

objects. In contrast, we consider a scenario where requests are handed over at peerings to the CDN. The papers closest to ours in the context of CDN-ISP collaboration are PaDIS [21] and FlowDirector [16]. Both systems collect topology and routing information from the ISP network to create a ranked mapping between the IPs of the ISP's users and the CDN's servers, e.g., based on distance or number of hops. The mapping is an additional input to the CDN's mapping system to improve content delivery by serving via servers at closer locations and shorter paths. Both aid the CDNs' mapping systems but only collect information from ISP's network and do not adapt it.

Optimization and resource allocation. The optimization problem considered in this paper is related to the virtual network embedding problem, which has been studied in the context of optical networks as well [78]–[83]. See [84] for an overview of existing solutions in this context. However, in contrast to the embedding problem, in our case, node allocations are already *given* (namely, the end-user locations and the hyper-giant's peering locations, modeled as super nodes), but the topology is *flexible*. The link-only embedding problem boils down to routing, but does not support the selection of multiple (content) locations. That said, the approaches in [78], [80], [82] for supporting addition and removal of IP links, could be reused in our setting; however, these approaches do not consider re-embedding of virtual links which maps to re-routing of demand in our scenario. We note that also works on mapping service function chains on optical networks, e.g., [85], are different from our approach as the optimization regards the mapping, not the topology.

VII. CONCLUSION

Motivated by emerging reconfigurable network infrastructures, this paper initiated the study of the potential benefits of dynamic and demand-aware re-optimizations. To this end, we presented an optimization framework to improve the delivery of *hyper-giant* traffic across ISP networks, leveraging short-term reconfigurability along three dimensions, topology, routing, and user mapping. Using extensive empirical analyses in collaboration with a large ISP, we found that such joint optimization is indeed worthwhile, and can benefit both the ISP and the *hyper-giant*. The observed benefits are two-fold. First, the required backbone capacity is reduced by 15% and second, path lengths are shortened by 30%, both on average and during the critical peak hour. Moreover, we show that infrastructure re-optimizations can also be leveraged in specific situations, for example during the COVID-19 pandemic or under link failures.

We understand our work as a first step, and believe that our approach opens several interesting directions for future research. In particular, while a more efficient use of the physical infrastructure benefits OPEX (e.g., in terms of fewer deployed lightpaths and reduced power consumption), it will be interesting to further evaluate the benefits for quality-of-experience of specific applications. Furthermore, while we have focused on a large ISP, it would also be interesting to study whether similar benefits can be obtained even in smaller ISPs, and at higher reconfiguration granularity.

ACKNOWLEDGMENTS

This work received funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - 316878574 and 438892507, and supported by the European Research Council (ERC), consolidator grant agreement 864228 (Adjust-Net).

REFERENCES

- [1] VNI Cisco. Cisco visual networking index: Forecast and trends, 2017–2022. *White Paper*, 1, 2018.
- [2] Sushant Jain, Alok Kumar, Subhasree Mandal, Joon Ong, Leon Poutievski, Arjun Singh, Subbaiah Venkata, Jim Wanderer, Junlan Zhou, Min Zhu, et al. B4: Experience with a globally-deployed software defined wan. *ACM CCR*, 43(4):3–14, 2013.
- [3] Chi-Yao Hong, Srikanth Kandula, Ratul Mahajan, Ming Zhang, Vijay Gill, Mohan Nanduri, and Roger Wattenhofer. Achieving high utilization with software-driven wan. *ACM CCR*, 43(4):15–26, 2013.
- [4] Parveen Patel, Deepak Bansal, Lihua Yuan, Ashwin Murthy, Albert Greenberg, David A Maltz, Randy Kern, Hemant Kumar, Marios Zikos, Hongyu Wu, et al. Ananta: Cloud scale load balancing. *ACM CCR*, 43(4):207–218, 2013.
- [5] Daniel E Eisenbud, Cheng Yi, Carlo Contavalli, Cody Smith, Roman Kononov, Eric Mann-Hielscher, Ardas Cililingiroglu, Bin Cheyney, Wentao Shang, and Jinnah Dylan Hosein. Maglev: A fast and reliable software network load balancer. In *Proc. USENIX NSDI 2016*, pages 523–535, 2016.
- [6] Pat Bosshart, Dan Daly, Glen Gibb, Martin Izzard, Nick McKeown, Jennifer Rexford, Cole Schlesinger, Dan Talayco, Amin Vahdat, George Varghese, et al. P4: Programming protocol-independent packet processors. *ACM CCR*, 44(3):87–95, 2014.
- [7] D. Lohar. Sonic: Software for open networking in the cloud. In <https://azure.github.io/SONiC/>, 2019.
- [8] Klaus-Tycho Foerster and Stefan Schmid. Survey of reconfigurable data center networks: Enablers, algorithms, complexity. *ACM SIGACT News*, 50(2):62–79, 2019.
- [9] Rachee Singh, Manya Ghobadi, Klaus-Tycho Foerster, Mark Filer, and Phillipa Gill. Radwan: rate adaptive wide area network. In *Proc. ACM SIGCOMM 2018*, pages 547–560, 2018.
- [10] Gagan Choudhury, David Lynch, Gaurav Thakur, and Simon Tse. Two Use Cases of Machine Learning for SDN-Enabled IP/Optical Networks: Traffic Matrix Prediction and Optical Path Performance Prediction [Invited]. *J. Opt. Commun. Netw.*, 10(10):D52, Oct 2018.
- [11] Simon Tse and Gagan Choudhury. Real-Time Traffic Management in AT&T's SDN-Enabled Core IP/Optical Network. In *Optical Fiber Communication Conference*, page Tu3H.2, San Diego, California, 2018. OSA.
- [12] N Sambo, A Giorgetti, A Sgambelluri, E Riccardi, F Fresi, F Cugini, and P Castoldi. Connection setup in open systems: Interfaces, amplification, and equalization. In *2018 European Conference on Optical Communication (ECOC)*, pages 1–3. IEEE, 2018.
- [13] Ori Gerstel, Clarence Filsfils, Thomas Telkamp, Matthias Gunkel, Martin Horneffer, Victor Lopez, and Arturo Mayoral. Multi-layer capacity planning for IP-optical networks. *IEEE Commun. Mag.*, 52(1):44–51, Jan 2014.
- [14] Luis Velasco, Alberto Castro, Daniel King, Ori Gerstel, Ramon Casellas, and Victor Lopez. In-operation network planning. *IEEE Commun. Mag.*, 52(1):52–60, Jan 2014.
- [15] Ciril Rozic, Marco Savi, Chris Matrakidis, Dimitrios Klonidis, and Domenico Siracusa. A Dynamic Multi-Layer Resource Allocation and Optimization Framework in Application-Centric Networks. *J. Lightwave Technol.*, 36(20):4908–4914, Oct 2018.
- [16] Enric Pujol, Ingmar Poesse, Johannes Zerwas, Georgios Smaragdakis, and Anja Feldmann. Steering hyper-giants' traffic at scale. In *Proc. ACM CoNEXT 2019*, page 82–95, 2019.
- [17] Coronavirus Disease (COVID-19) Situation Reports. <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports>.
- [18] C. Labovitz. Internet Traffic 2009-2019. APRICOT 2019, 2019.
- [19] F. Wohlfart, N. Chatzis, C. Dabanoglu, G. Carle, and W. Willinger. Leveraging Interconnections for Performance: The Serving Infrastructure of a Large CDN. In *Proc. ACM SIGCOMM 2018*, pages 206–220, 2018.

- [20] V. Giotsas, M. Luckie, B. Huffaker, and K. Claffy. Inferring Complex AS Relationships. In *Proc. ACM IMC 2014*, pages 23–30, 2014.
- [21] Ingmar Poese, Benjamin Frank, Bernhard Ager, Georgios Smaragdakis, and Anja Feldmann. Improving content delivery using provider-aided distance information. In *Proc. ACM IMC 2010*, pages 22–34, 2010.
- [22] Benjamin Frank, Ingmar Poese, Georgios Smaragdakis, Steve Uhlig, and Anja Feldmann. Content-aware Traffic Engineering. *arXiv:1202.1464 [cs]*, Feb 2012.
- [23] Gregor Maier, Anja Feldmann, Vern Paxson, and Mark Allman. On dominant characteristics of residential broadband internet traffic. In *Proc. ACM IMC '09*, pages 90–102, 2009.
- [24] Jeffrey Ertman and K.K. Ramakrishnan. Understanding the super-sized traffic of the super bowl. In *Proc. ACM IMC 2013*, pages 353–360, 2013.
- [25] Jeremias Blendin, Fabrice Bendfeldt, Ingmar Poese, Boris Koldehofe, and Oliver Hohlfeld. Dissecting Apple's Meta-CDN during an iOS Update. In *Proc. ACM IMC 2018*, pages 408–414, 2018.
- [26] DE-CIX. Big upswing in internet usage due to covid-19 measures. <https://www.de-cix.net/en/news-events/news/big-upswing-in-internet-usage-due-to-covid-19-measures>. Accessed: 2020-06-22.
- [27] Ward Van Heddeghem, Bart Lannoo, Didier Colle, Mario Pickavet, and Piet Demeester. A quantitative survey of the power saving potential in ip-over-wdm backbone networks. *IEEE Communications Surveys & Tutorials*, 18(1):706–731, 2014.
- [28] Martin Birk, Gagan Choudhury, Bruce Cortez, Alvin Goddard, Narayan Padi, Aswatnarayan Raghuram, Kathy Tse, Simon Tse, Andrew Wallace, and Kang Xi. Evolving to an SDN-enabled isp backbone: Key technologies and applications. *IEEE Commun. Mag.*, 54(10):129–135, Oct 2016.
- [29] JP. Vasseur and JL. Le Roux. Path computation element (pce) communication protocol (pcep). RFC 5440, RFC Editor, March 2009. <http://www.rfc-editor.org/rfc/rfc5440.txt>.
- [30] The state of traffic engineering - an ISP's perspective. <https://blog.apnic.net/2018/08/24/the-state-of-traffic-engineering-an-isp-perspective/>.
- [31] C. Filsfils, S. Previdi, L. Ginsberg, B. Decraene, S. Litkowski, and R. Shakir. Segment routing architecture. RFC 8402, RFC Editor, July 2018. <http://www.rfc-editor.org/rfc/rfc8402.txt>.
- [32] E. Crabbe, I. Minei, J. Medved, and R. Varga. Path computation element communication protocol (pcep) extensions for stateful pce. RFC 8231, RFC Editor, September 2017. <http://www.rfc-editor.org/rfc/rfc8231.txt>.
- [33] Fangfei Chen, Ramesh K Sitaraman, and Marcelo Torres. End-user mapping: Next generation request routing for content delivery. *ACM CCR*, 45(4):167–181, 2015.
- [34] K-K. Yap, M. Motiwala, J. Rahe, S. Padgett, M. Holliman, G. Baldus, M. Hines, T. Kim, A. Narayanan, A. Jain, V. Lin, C. Rice, B. Rogan, A. Singh, B. Tanaka, M. Verma, P. Sood, M. Tariq, M. Tierney, D. Trumic, V. Valancius, C. Ying, M. Kallahalla, B. Koley, and A. Vahdat. Taking the Edge off with Espresso: Scale, Reliability and Programmability for Global Internet Peering. In *Proc. ACM SIGCOMM 2017*, pages 432–445, 2017.
- [35] B. Schlinder, H. Kim, T. Cui, E. Katz-Bassett, H. V. Madhyastha, I. Cunha, J. Quinn, S. Hasan, P. Lapukhov, and H. Zeng. Engineering Egress with Edge Fabric: Steering Oceans of Content to the World. In *Proc. ACM SIGCOMM 2017*, pages 418–431, 2017.
- [36] B. M. Maggs and R. K. Sitaraman. Algorithmic Nuggets in Content Delivery. *ACM CCR*, 45(3):52–66, 2015.
- [37] R. Alimi, R. Penno, Y. Yang, S. Kiesel, S. Previdi, W. Roome, S. Shalunov, and R. Woundy. Application-layer traffic optimization (alto) protocol. RFC 7285, RFC Editor, September 2014. <http://www.rfc-editor.org/rfc/rfc7285.txt>.
- [38] Giovane CM Moura, John Heidemann, Ricardo de O Schmidt, and Wes Hardaker. Cache me if you can: effects of dns time-to-live. In *Proc. ACM IMC 2019*, pages 101–115, 2019.
- [39] Krzysztof Walkowiak and Mirosław Klinkowski. Joint anycast and unicast routing for elastic optical networks: Modeling and optimization. In *Proc. IEEE ICC 2013*, pages 3909–3914, 2013.
- [40] IBM ILOG CPLEX Optimization Studio 12.9. <https://www.ibm.com/products/ilog-cplex-optimization-studio>.
- [41] Thomas Favale, Francesca Soro, Martino Trevisan, Idilio Drago, and Marco Mellia. Campus traffic and e-learning during covid-19 pandemic. *Computer Networks*, page 107290, 2020.
- [42] Anja Feldmann, Oliver Gasser, Franziska Lichtblau, Enric Pujol, Ingmar Poese, Christoph Dietzel, Daniel Wagner, Matthias Wichtlhuber, Vallina-Rodrigue Narseo Tapiador, Juan and, Oliver Hohlfeld, and Georgios Smaragdakis. The Lockdown Effect: Implications of the COVID-19 Pandemic on Internet Traffic. In *ACM IMC 2020*, pages 1–18, 2020.
- [43] BBC. Netflix to cut streaming quality in europe for 30 days. <https://www.bbc.com/news/technology-51968302>. Accessed: 2020-06-22.
- [44] Sandvine. Over 43% of the internet is consumed by Netflix, Google, Amazon, Facebook, Microsoft, and Apple: Global Internet Phenomena Spotlight. <https://www.sandvine.com/blog/netflix-vs.-google-vs.-amazon-vs.-facebook-vs.-microsoft-vs.-apple-traffic-share-of-internet-brands-global-internet-phenomena-spotlight>.
- [45] Tom Leighton. Improving performance on the internet. *Communications of the ACM*, 52(2):44–51, 2009.
- [46] Tobias Flach, Nandita Dukkupati, Andreas Terzis, Barath Raghavan, Neal Cardwell, Yuchung Cheng, Ankur Jain, Shuai Hao, Ethan Katz-Bassett, and Ramesh Govindan. Reducing web latency: the virtue of gentle aggression. *ACM CCR*, 43(4):159–170, 2013.
- [47] Wenjie Jiang, Rui Zhang-Shen, Jennifer Rexford, and Mung Chiang. Cooperative content distribution and traffic engineering in an ISP network. In *Proc. ACM SIGMETRICS '09*, pages 239–250, 2009.
- [48] Nicolas Herbaut, Daniel Negru, Yiping Chen, Pantelis A. Frangoudis, and Adlen Ksentini. Content Delivery Networks as a Virtual Network Function: A Win-Win ISP-CDN Collaboration. In *IEEE GLOBECOM 2016*, pages 1–6, Dec 2016.
- [49] Matthias Wichtlhuber, Robert Reinecke, and David Hausheer. An SDN-Based CDN/ISP Collaboration Architecture for Managing High-Volume Flows. *IEEE Trans. Netw. Serv. Manage.*, 12(1):48–60, Mar 2015.
- [50] Matthias Wichtlhuber, Jan Kessler, Sebastian Buckner, Ingmar Poese, Jeremias Blendin, Christian Koch, and David Hausheer. SoDA: Enabling CDN-ISP collaboration with software defined anycast. In *Proc. IFIP Networking 2017*, pages 1–9, 2017.
- [51] Google Fiber Inc. <https://fiber.google.com/>.
- [52] Amazon receives FCC approval for Project Kuiper satellite constellation. <https://blog.aboutamazon.com/company-news/amazon-receives-fcc-approval-for-project-kuiper-satellite-constellation>.
- [53] ATT TV. <https://www.att.com/tv/>.
- [54] Xfinity Digital Cable TV. <https://www.xfinity.com/learn/digital-cable-tv>.
- [55] Brigitte Jaumard, Huy Quang Duong, Romualdas Armonavicius, Todd Morris, and Petar Djukic. Efficient real-time scalable make-before-break network re-routing. *IEEE/OSA Journal of Optical Communications and Networking*, 11(3):52–66, 2019.
- [56] Ramon Casellas, Ricardo Martinez, Ricard Vilalta, and Raul Munoz. Control, Management, and Orchestration of Optical Networks: Evolution, Trends, and Challenges. *J. Lightwave Technol.*, 36(7):1390–1402, Apr 2018.
- [57] Jing Wu. A survey of WDM network reconfiguration: Strategies and triggering methods. *Computer Networks*, 55(11):2622–2645, Aug 2011.
- [58] Mohit Chamania, Marcel Caria, and Admela Jukan. Achieving IP routing stability with optical bypass. In *2009 IEEE 3rd International Symposium on Advanced Networks and Telecommunication Systems (ANTS)*, pages 1–3, 2009.
- [59] Mohit Chamania and Admela Jukan. A Comparative Analysis of the Effects of Dynamic Optical Circuit Provisioning on IP Routing. *IEEE/ACM Trans. Networking*, 22(2):429–442, Apr 2014.
- [60] B. Ramamurthy and A. Ramakrishnan. Virtual topology reconfiguration of wavelength-routed optical WDM networks. In *IEEE Globecom '00. Conference Record (Cat. No.00CH37137)*, volume 2, pages 1269–1275, San Francisco, CA, USA, 2000.
- [61] P. Papanikolaou, K. Christodouloulopoulos, and E. Varvarigos. Incremental planning of multi-layer elastic optical networks. In *Proc. International Conference on Optical Network Design and Modeling (ONDM) 2017*, pages 1–6, May 2017.
- [62] Mirosław Klinkowski and Krzysztof Walkowiak. On the advantages of elastic optical networks for provisioning of cloud computing traffic. *IEEE Network*, 27(6):44–51, Nov 2013.
- [63] Ioannis Tomkos, Ćiril Roĳić, Marco Savi, Pontus Sköldström, Victor Lopez, Mohit Chamania, Domenico Siracusa, Chris Matrakidis, Dimitrios Klonidis, and Ori Gerstel. Application aware multilayer control and optimization of elastic WDM switched optical networks. In *Optical Fiber Communication Conference*, page Th1D.4, San Diego, California, 2018. OSA.
- [64] Satyajeet Singh Ahuja, Varun Gupta, Vinayak Dangu, Soshant Bali, Abishek Gopalan, Hao Zhong, Petr Lapukhov, Yiting Xia, and Ying Zhang. Capacity-efficient and uncertainty-resilient backbone network planning with hose. In *Proc. ACM SIGCOMM 2021*, page 547–559, 2021.
- [65] Srivatsan Balasubramanian, Satyajeet Ahuja, Gaya Nagarajan, Andrea Celletti, and Frantisek Foston. Multilayer planning for Facebook scale worldwide network. In *2017 International Conference on Optical*

Network Design and Modeling (ONDM), pages 1–6, Budapest, May 2017. IEEE.

- [66] Jordi Perelló, Krzysztof Walkowiak, Mirosław Klinkowski, Salvatore Spadaro, and Davide Careglio. Joint content placement and lightpath routing and spectrum assignment in CDNs over elastic optical network scenarios. *Computer Communications*, 77:72–84, Mar 2016.
- [67] Xin Jin, Yiran Li, Da Wei, Siming Li, Jie Gao, Lei Xu, Guangzhi Li, Wei Xu, and Jennifer Rexford. Optimizing bulk transfers with software-defined optical wan. In *Proc. ACM SIGCOMM 2016*, pages 87–100. ACM, 2016.
- [68] Monia Ghobadi, Ratul Mahajan, Amar Phanishayee, Nikhil Devanur, Janardhan Kulkarni, Gireeja Ranade, Pierre-Alexandre Blanche, Houman Rastegarfar, Madeleine Glick, and Daniel Kilper. Projector: Agile reconfigurable data center interconnect. In *Proc. ACM SIGCOMM 2016*, pages 216–229, 2016.
- [69] Mowei Wang, Yong Cui, Shihan Xiao, Xin Wang, Dan Yang, Kai Chen, and Jun Zhu. Neural network meets dcn: Traffic-driven topology adaptation with deep learning. *Proceedings of the ACM on Measurement and Analysis of Computing Systems*, 2(2):26, 2018.
- [70] Shaileshh Bojja Venkatakrishnan, Mohammad Alizadeh, and Pramod Viswanath. Costly circuits, submodular schedules and approximate carathéodory theorems. In *ACM SIGMETRICS Performance Evaluation Review*, volume 44, pages 75–88. ACM, 2016.
- [71] Klaus-Tycho Foerster, Monia Ghobadi, and Stefan Schmid. Characterizing the algorithmic complexity of reconfigurable data center architectures. In *Proc. ACM/IEEE Symposium on Architectures for Networking and Communications Systems (ANCS)*, 2018.
- [72] Thomas Fenz, Klaus-Tycho Foerster, Stefan Schmid, and Anais Villedieu. Efficient non-segregated routing for reconfigurable demand-aware networks. In *Proc. IFIP Networking*, 2019.
- [73] Arjun Roy, Hongyi Zeng, Jasmeet Bagga, George Porter, and Alex C Snoeren. Inside the social network’s (datacenter) network. 45(4):123–137, 2015.
- [74] Hongqiang Harry Liu, Raajay Viswanathan, Matt Calder, Aditya Akella, Ratul Mahajan, Jitendra Padhye, and Ming Zhang. Efficiently Delivering Online Services over Integrated Infrastructure. In *Proc. USENIX NSDI 2016*, pages 77–90, 2016.
- [75] Mostafa Ammar, Ellen Zegura, and Yimeng Zhao. A Vision for Zero-Hop Networking (ZeN). In *2017 IEEE 37th International Conference on Distributed Computing Systems (ICDCS)*, pages 1765–1770, Atlanta, GA, USA, Jun 2017.
- [76] Abhigyan Sharma, Arun Venkataramani, and Ramesh K Sitaraman. Distributing Content Simplifies ISP Traffic Engineering. In *Proc. ACM SIGMETRICS 2013*, page 14, 2013.
- [77] Khaled Diab and Mohamed Hefeeda. Joint Content Distribution and Traffic Engineering of Adaptive Videos in Telco-CDNs. In *Proc. IEEE INFOCOM 2019*, pages 1342–1350, Apr 2019.
- [78] Sunny Shakya, Nabina Pradhan, Xiaojun Cao, Zilong Ye, and Chunming Qiao. Virtual network embedding and reconfiguration in elastic optical networks. In *Proc. IEEE GLOBECOM 2014*, pages 2160–2165, Austin, TX, USA, Dec 2014.
- [79] Leonard Nonde, Taisir E. H. El-Gorashi, and Jaafar M. H. Elmirghani. Energy Efficient Virtual Network Embedding for Cloud Networks. *J. Lightwave Technol.*, 33(9):1828–1849, May 2015.
- [80] Jiawei Zhang, Yuefeng Ji, Mei Song, Hui Li, Rentao Gu, Yongli Zhao, and Jie Zhang. Dynamic Virtual Network Embedding Over Multilayer Optical Networks. *J. Opt. Commun. Netw.*, 7(9):918, Sep 2015.
- [81] Raouf Boutaba, Nashid Shahriar, and Siavash Fathi. Elastic Optical Networking for 5G Transport. *J Netw Syst Manage*, 25(4):819–847, Oct 2017.
- [82] Shihabur Rahman Chowdhury, Sara Ayoubi, Reaz Ahmed, Nashid Shahriar, Raouf Boutaba, Jeebak Mitra, and Liu Liu. Multi-Layer Virtual Network Embedding. *IEEE Trans. Netw. Serv. Manage.*, 15(3):1132–1145, Sep 2018.
- [83] Francisco Carpio, Samia Dhahri, and Admela Jukan. VNF placement with replication for Load balancing in NFV networks. In *Proc. IEEE ICC 2017*, pages 1–6, Paris, France, May 2017.
- [84] Danilo Bórquez-Paredes, Alejandra Beghelli, and Ariel Leiva. Network Virtualization Over Elastic Optical Networks: A Survey of Allocation Algorithms. In Rastislav Roka, editor, *Optical Fiber and Wireless Communications*. InTech, Jun 2017.
- [85] Zilong Ye, Xiaojun Cao, Jianping Wang, Hongfang Yu, and Chunming Qiao. Joint topology design and mapping of service function chains for efficient, scalable, and reliable network functions virtualization. *IEEE Network*, 30(3):81–87, May 2016.



Johannes Zerwas received his M.Sc. degree in Electrical Engineering and Information Technology from Technical University of Munich (TUM), Germany, in 2018. He joined the Chair of Communication Networks at the TUM as a research and teaching associate in February 2018. His research focuses on reconfigurable topologies and data-driven networking algorithms.

Ingmar Poesse is ...

ipoesse.jpg



Stefan Schmid is a Professor at the Technical University of Berlin, Germany, working part-time for the Fraunhofer Institute for Secure Information Technology (SIT) in Germany as well as for the Faculty of Computer Science, University of Vienna in Austria. MSc and PhD at ETH Zurich, Postdoc at TU Munich and University of Paderborn, Senior Research Scientist at T-Labs in Berlin, Associate Professor at Aalborg University, Denmark, and Full Professor at the University of Vienna, Austria. Stefan Schmid received the IEEE Communications Society ITC Early Career Award 2016 and an ERC Consolidator Grant 2019.



network measurements and simulations, and the application of machine learning and artificial intelligence in communication networks.

Andreas Blenk received the Dr.-Ing. degree (Ph.D.) from the Technical University of Munich in 2018 with the highest distinction (summa cum laude). He joined the Chair of Communication Networks at the Technical University of Munich in June 2012, where he is currently working as a senior researcher and associate lecturer. Since March 2019, he has also been a senior research fellow at the Communication Technologies group of the Faculty of Computer Science of the University of Vienna. In his research, he focuses on self-driving, programmable networks,