Deferring CAPEX Investments in Multi-Layer Networks through IP Traffic-Dependent Expansion Stages

Michael Duelli, Anke Endler, and Michael Menth
University of Würzburg, Institute of Computer Science, Chair of Distributed Systems
Am Hubland, 97074 Würzburg, Germany
E-mail: {duelli|anke.endler|menth}@informatik.uni-wuerzburg.de

Abstract

Network design is performed for a traffic load in the future. This traffic load is far greater than traffic loads observed at intermediate stages. In this paper, we propose to plan a multi-layer network for such a future traffic load, but instead of provisioning the full network equipment at once, we suggest to defer the deployment of expensive IP interfaces and other enabling equipment only to stages when they are really needed to carry traffic. We develop an algorithm that concentrates IP traffic at an intermediate stage on a subset of those IP interfaces that are required to carry the full load in the future. This reduces the set of used IP interfaces which need to be installed at intermediate stages, and CAPEX investments for the other equipment can be deferred to the future.

1 Introduction

All of today's traffic forecasts predict an ongoing increase of IP data traffic which is said to be at least linear or even exponential in the considered forecast period. As a consequence, network providers have to continually upgrade their network configurations to keep up with increasing traffic requirements and prevent violations of *service level agreements* (SLA) for established services. Since the revenue of the providers on a per-bit basis decreases at the same time, network providers have to take special care of network upgrades and related investment cost, i.e. *capital expenditure* (CAPEX), to ensure profitability.

As traffic forecasts incorporate a certain amount of uncertainty regarding the expected traffic volume and its arrival time, network providers have to over-dimension their networks in order to guarantee that all future traffic requirements can be handled. Since over-dimensioning is a direct trade-off to cost minimality, network providers want to plan their networks for several years in advance such that upgrades for increasing traffic requirements can be incrementally realized and integrated into the existing network infrastructure.

In this paper, we consider the CAPEX-aware design of multi-layer networks incorporating *optical transport network* (OTN) and IP/MPLS technology for given traffic requirements. We present an algorithm for the design of IP topologies which are suited for given traffic requirements and can be extended up to a previously defined expansion stage which fulfills future traffic requirements. We evaluate the algorithm upon two German reference topologies for a homogeneous and a population-based traffic pattern with increasing traffic requirements.

The remainder of this paper is structured as follows. Section 2 introduces the problem formulation and gives an overview of related work. Section 3 explains the models used for CAPEX and multi-layer technology as well as traffic and resources. In Section 4, an algorithm is presented for IP topology design with given traffic forecasts which is evaluated in Section 5. Section 6 summarizes this paper.

2 Problem Formulation and Related Work

2.1 Problem Formulation

Traffic forecasts like [1] predict a *compound annual growth rate* (CAGR) of 40% for overall IP data traffic volume between 2008 and 2013. Such forecasts are updated on a yearly basis to keep up with the real traffic development. Since a radical change of network configurations in a short time bears high risks and costs, network providers heavily rely on such forecasts in order to be able to deal with future traffic requirements. Moreover, traffic usually does not grow linearly but rather exponentially so that most of the transmission capabilities of a designed network are needed only in the future.

Traffic forecasts incorporate several uncertainties regarding the expected overall traffic volume, the traffic between certain sites, and the timely occurrence of traffic flows in particular. As a consequence, network providers have to over-dimension their networks to a suitable extend and keep a certain degree of flexibility concerning the upgrade of network configurations to deal with anomalies of the traffic forecast.

To enable flexible network upgrades and to decrease the uncertain anomalies of short-term forecasts within

The authors have been funded by the Federal Ministry of Education and Research of the Federal Republic of Germany (BMBF Förderkennzeichen 01BP0775). Their work is part of the EUREKA project "100 Gbit/s Carrier-Grade Ethernet Transport Technologies (CELTIC CP4-001)". The authors alone are responsible for the content of the paper.

the design of a multi-layer network, we presume that a target expansion stage for a network configuration is planned for a future point in time, e.g. a couple of years from now. This target expansion stage is able to carry a given set of future traffic requirements. We assume traffic forecasts for earlier points in time whose requirements are lower and develop an algorithm that calculates a subset of the future target network configuration that suits to the given traffic forecasts. The reduced network configurations allow straightforward upgrades to the target expansion stage.

2.2 Related Work

Traffic engineering in IP/MPLS networks has been investigated for more than a decade. The main concerns of existing publications are blocking probabilities as well as resource utilization with or without considering resilience. The authors of [2] perform a mathematical approach for on-line routing of bandwidth guaranteed *label switched paths* (LSP) with restoration. More recent publications presume the availability of *generalized MPLS* (GMPLS) to consider traffic engineering on GMPLS networks with different transport networks as done in [3]. Especially, IP services over OTN networks have been considered. For instance, the authors of [4] present two algorithms for dynamic provisioning of SLA-guaranteed IP services in arbitrary IP-over-optical networks, while the authors of [5] present dynamic routing algorithms for transparent optical networks. All evaluations of dynamic traffic routing are performed by means of simulation.

In previous work, we have analyzed the impact of an underlying optical transport network with different grooming methodologies on IP traffic in case of failures [6] and presented a multi-layer network design algorithm [7]. In [8], an overview of requirements and interworking of network design and dynamic provisioning mechanisms was given whose results lead to the combination of network design and dynamic provisioning. To the best of our knowledge, no heuristics have been published that create multi-layer network topologies which are suited for incremental equipment upgrades to keep up with increasing traffic requirements to a certain planning extent.

3 Modeling

In this section, models for network equipment and *capital expenditure* (CAPEX), technology interconnection, traffic and resources are described.

3.1 Network Equipment and CAPEX Model

A network equipment model defines the available technologies and their possible interconnections. In turn, a CAPEX model associates cost values with each networking equipment. In [9], a network equipment and CAPEX model was given covering IP/MPLS, Ethernet, SDH/SONET, and *optical transport network* (OTN) technologies. The cost values are vendor-independent and normalized to the cost of a 10 Gbit/s WDM transponder. [10] lists the available network equipment and corresponding CAPEX values for this model.

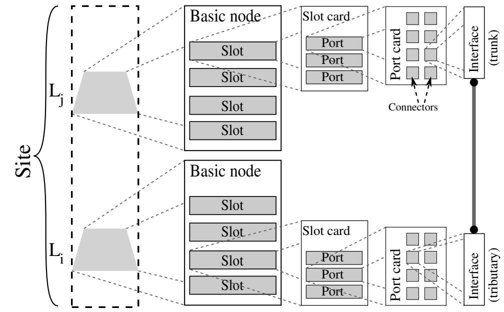


Figure 1: Modular set-up and interconnection of two technology layers L_i and L_j within a single site using modular components.

Networking equipment is built in a modular fashion to simplify exchange of components, increase flexibility, and keep CAPEX low. We strictly split the equipment of all technologies considered in [10] into four component groups. *Basic nodes* provide core functionality within a technology, like power supply, cooling, and backbone switching for all incorporated components. *Slot cards* can be plugged into each slot of a basic node and provide access to the backbone switching of the basic node for *port cards* which can be plugged into each port of a slot card, in turn. The port cards can be populated with *interfaces* which send or receive data using a certain encoding. An exemplary assembly of the four component groups within a site is illustrated in Fig. 1.

As described in [7], we strictly applied the four component model on all technologies in contrast to the original model. Consequently, we provided zero-cost dummy components in case a technology model was lacking a component group. Furthermore, we separate the OTN technology into an *optical channel* (OCh) layer dealing with wavelengths and lightpaths as well as an *optical multiplex section* (OMS) layer handling and switching bundles of wavelengths on fibers.

3.2 Technology Interconnection Model

Within a single technology at a single site, the assembly of the four component groups is only restricted by equipment constraints. For instance, the number of slots or an upper limit for the sum of requested switching capacities cannot be exceeded. The interconnection of different technologies within a site or across remote sites requires that the transmitting interface encapsulates the data in a format that can be processed by the receiving interface. Hence, a pair of transmitting and receiving interfaces have to be *compatible*.

According to the communication direction in the layer

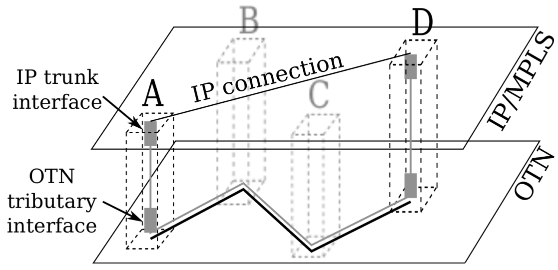


Figure 2: Interfaces for an IP connection.

hierarchy, these interfaces are called *trunk* if they communicate downwards in the layer hierarchy or *tributary* if they communicate upwards, as depicted in Fig. 1. Depending on their technology and ability, interfaces use a specific modulation and framing to en/decapsulate data. The framed data is transmitted on an interface specific bit rate, the interface's *capacity*. Two interfaces must use the same encoding of data, i.e. modulation, framing, and capacity, to be compatible.

An interconnection of two sites across multiple layers and intermediate sites may involve repeated encapsulation and aggregation of data at each technology layer as well as the decapsulation and deaggregation of data at intermediate sites. Fig. 2 illustrates an IP transmission along four sites on which data is en- and decapsulated at the source and destination sites to be able to process data at the responsible layers.

3.3 Traffic Model

We consider IP/MPLS technology which uses an explicit *label switched path* (LSP) for each traffic demand and thereby enables arbitrary IP traffic engineering. To be able to neglect the resource utilization on the OTN layer, we presume that all demands have a fixed bandwidth requirement c_d which corresponds to an *optical data unit* (ODU) which fits into the OTN containers.

We control the traffic load in the network by a parameter d_{tot} which defines the overall number of IP demands that is to be routed between the sites $\mathcal{V}$ of a network. We consider a homogeneous and a population-based traffic model.

With the homogeneous traffic model, there are d_{hom} directed IP demands between each pair of sites $x, y \in \mathcal{V}$. Hence, we have

$$d_{xy} = d_{\text{hom}} \tag{1}$$

and

$$d_{\text{tot}} = d_{\text{hom}} \cdot |\mathcal{V}| \cdot (|\mathcal{V}| - 1). \tag{2}$$

For the population-based traffic model [11], the amount of traffic between two sites $x, y \in \mathcal{V}$ is determined by the ratio of population of these two sites to all pairs of sites in the network. We define the number of undirected IP traffic demands between two sites $x, y \in \mathcal{V}$ by

$$d_{x,y} = d_{\text{tot}} \cdot \frac{\pi(x) \cdot \pi(y)}{\sum_{a,b \in \mathcal{V}, a \neq b} \pi(a) \cdot \pi(b)} \tag{3}$$

where $\pi(x)$ is the population of site $x \in \mathcal{V}$. As we consider directed demands, the overall traffic between sites x and y is split proportionally into both directions via

$$d'_{xy} = d_{x,y} \cdot \frac{\pi(x)}{\pi(x) + \pi(y)}. \tag{4}$$

In general, d'_{xy} cannot be supposed to be an integer value, so we use $d_{xy} = \lceil d'_{xy} \rceil$ instead.

3.4 Resource Model

All links in the multi-layer network are directed. Each edge in the IP/MPLS topology is realized in the underlying OTN using trunk and tributary interfaces as illustrated in Fig. 2. The IP trunk interface that connects a link l in the IP topology to the OTN layer provides a certain capacity $c_{\text{IP}}(l)$. The spare capacity of IP link l is denoted by $c_{\text{spare}}(l)$. The ratio of $c_{\text{spare}}(l)$ and $c_{\text{IP}}(l)$ marks the utilization of l.

4 Algorithm

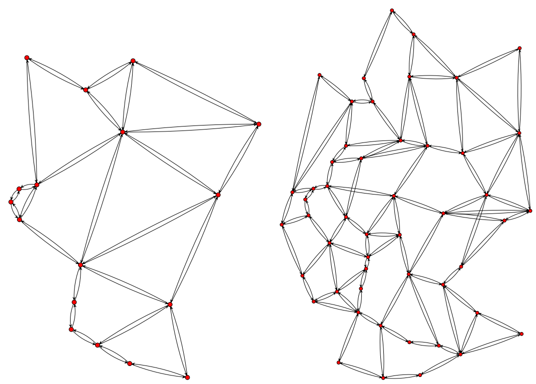
We assume an IP topology $G_{\text{IP}} = (\mathcal{V}, \mathcal{E}_{\text{IP}})$ where $\mathcal{V}$ are the sites of the network and $\mathcal{E}_{\text{IP}} \subseteq \mathcal{V} \times \mathcal{V}$ are the IP links. This IP topology is assumed to be given by the design of an underlying multi-layer network.

The IP topology G_{IP} is designed for a future traffic matrix $\mathcal{D}_{\text{fut}}$, i.e. a sufficient amount of resources is provided during network design to route all demands in $\mathcal{D}_{\text{fut}}$. We present an algorithm that is able to yield multi-layer network configurations and IP network topologies that suffice for given IP traffic requirements and can be incrementally extended to the final network expansion stage of $\mathcal{D}_{\text{fut}}$. Such a given IP traffic requirement is represented by a traffic matrix $\mathcal{D}_{\text{cur}}$ which defines a number of demands d_{xy} between two sites $x, y \in \mathcal{V}$ as described in Section 3.3.

Algorithm 1 processes all links in the IP topology which are in a candidate list $L = \mathcal{E}_{\text{IP}}$. The algorithm selects a link $l = (x, y) \in L$ from the candidate list according to the probability $p(l)$ defined by the ratio of spare capacity of link l to the overall spare capacity of the remaining candidate links. In case, no link in the candidate list L has left any spare capacity, any of the

Algorithm 1 Link reduction algorithm.

$L = \mathcal{E}_{\text{IP}}$ {Initialize with IP topology}
while $L \neq \emptyset$ **do**
  Select link $l \in L$ using Equation (5)
  $L = L \setminus \{l\}$
  if All traffic on l can be rerouted **then**
    Reroute traffic carried by l
    $\mathcal{E}_{\text{IP}} = \mathcal{E}_{\text{IP}} \setminus \{l\}$ {Remove l from IP topology}
  end if
end while



(a) Nobel Germany (G17) ($|\mathcal{V}| = 17, |\mathcal{E}_0| = 52$)   (b) Germany50 (G50) ($|\mathcal{V}| = 50, |\mathcal{E}_0| = 176$)

Figure 3: The two considered physical topologies from [12] with directed edges.

edges in the candidate list is uniformly chosen. The function $p : L \mapsto [0,1] \subseteq \mathbb{R}$ calculates the described link candidate selection probability and is defined as

$$p(l) = \begin{cases} \dfrac{c_{\text{spare}}(l)}{\sum_{l' \in L} c_{\text{spare}}(l')}, & \sum_{l' \in L} c_{\text{spare}}(l') \neq 0 \\ \dfrac{1}{|L|} & \text{otherwise.} \end{cases} \quad (5)$$

The algorithm checks whether the traffic on link l can be rerouted on the remaining IP topology. Therefore, we use a weight function $w : \mathcal{E}_{\text{IP}} \mapsto \mathbb{R}^+$ which is defined as

$$w(l) = 1 + w_{\max} \cdot \frac{c_{\text{spare}}(l)}{c_{\text{IP}}(l)}, \quad (6)$$

where $w_{\max} \in \mathbb{R}_0^+$ is a control parameter and $\frac{c_{\text{spare}}(l)}{c_{\text{IP}}(l)}$ defines the current resource utilization of link $l \in \mathcal{E}_{\text{IP}}$. The weight function w is applied on the IP topology $\mathcal{E}_{\text{IP}}$ and a shortest path algorithm is used to find alternative routings for the traffic on link l. Thus, link l is kept unused by preferring the existing resources of the IP topology and increasing their utilization. The lower the utilization of a link, the lower is its associated weight which makes it more likely to be used.

The intended effect of our algorithm is to defer the utilization of as many IP links as possible. This is possible as long as sufficient spare resources exist in the network to reroute the link's traffic on alternative paths. Hereby, the cost for this link and its related equipment across the multi-layer network can be saved until the traffic load reaches a certain threshold.

5 Evaluation

Our algorithm requires an IP topology provided by an underlying multi-layer network and an IP traffic pattern.

The evaluations are performed on two German reference topologies taken from [12] which are used as physical fiber topologies and have been equiped with directed links: the *Nobel Germany* topology (G17) depicted in Fig. 3(a) and the *Germany50* topology (G50) depicted in Fig. 3(b) incorporating all sites of G17. Upon both topologies, we design a multi-layer network incorporating OTN and IP technology using the *auxiliary cross layer* (AXL) algorithm presented in [7]. The resulting multi-layer networks are suited to carry a future traffic matrix $\mathcal{D}_{\text{fut}}$ defining d_{xy}^{fut} IP traffic demands from site x to y. Each demand in $\mathcal{D}_{\text{fut}}$ has a bandwidth requirement of $c_d = 1$ Gbit/s which corresponds to the OTN container ODU-0 as described in Section 3.3. The multi-layer network underlying the IP topology carries these IP traffic demands by providing a total of $\lceil d_{xy}^{\text{fut}} \cdot c_d / c_{\text{IP}} \rceil$ parallel IP links from site x to y with capacity c_{IP}. For evaluation, we use $c_{\text{IP}} = 10$ Gbit/s and denoted the resulting IP traffic matrix by $\mathcal{D}'_{\text{fut}}$.

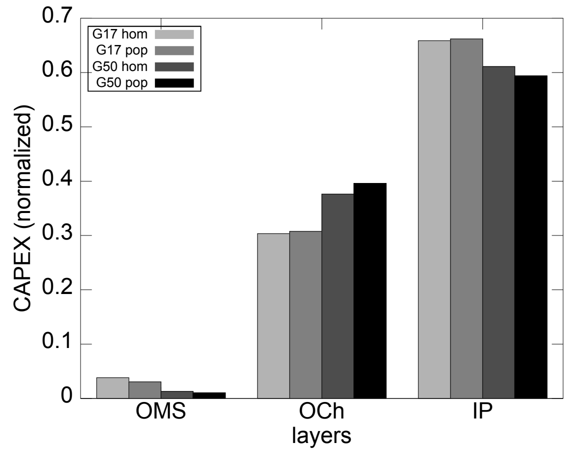


Figure 4: CAPEX split by OMS, OCh, and IP layers for the considered scenarios.

Upon both topologies, we apply the two traffic models described in Section 3.3. On the one hand, we use a homogeneous traffic pattern which allows to fully utilize the existing resources. On the other hand, we consider a population-based traffic model [11] which uses population data to determine the number of data flows between two sites and results in a heterogeneous traffic distribution.

The parameters of the resulting multi-layer networks and corresponding $\mathcal{D}'_{\text{fut}}$ are listed in Table 1. Fig. 4 illustrates the CAPEX of the four scenarios split by *op-*

Table 1: Parameters of the multi-layer networks designed for $\mathcal{D}'_{\text{fut}}$.

Topology & Traffic	$\lvert\mathcal{D}'_{\text{fut}}\rvert$	CAPEX by design for $\mathcal{D}'_{\text{fut}}$
G17 homogeneous	272	4595.76
G17 population	330	5752.14
G50 homogeneous	1920	50406.85
G50 population	3146	64438.07

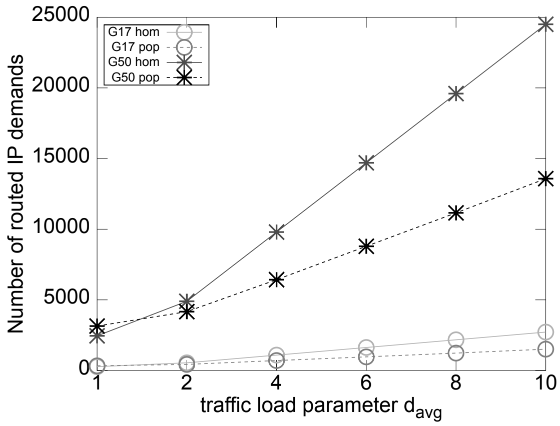


Figure 5: Overall number of IP traffic demands.

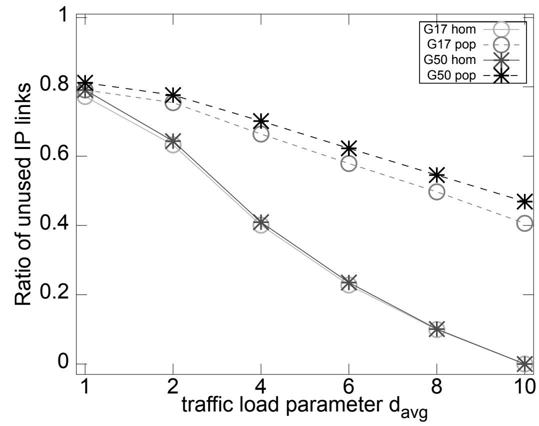


Figure 6: Evaluation of unused IP links for different traffic loads d_{avg}.

tical multiplex section (OMS), *optical channel* (OCh) and *Internet protocol* (IP) technology. Fig. 4 shows that IP technology constitutes the highest cost factor ($> 50\%$) in all considered multi-layer networks while 30% to 40% go to the OCh equipment which contains the optical transponders and sets up lightpaths in the OTN layer. The remaining CAPEX goes to the OMS equipment like *optical add-drop multiplexers* (OADM) and *optical cross connects* (OXC) which switch the lightpaths in the OTN layer.

From such a designed multi-layer network, our algorithm extracts a reduced IP topology for a given IP traffic matrix $\mathcal{D}_{\text{cur}}$. While the initial planning traffic matrix $\mathcal{D}_{\text{fut}}$ corresponds to the traffic that is expected to happen in the final expansion stage of the network, expansion stages of the multi-layer network are to be found that (a) fit to the currently considered traffic matrix $\mathcal{D}_{\text{cur}}$, e.g. at an earlier point in time, and (b) can be incrementally expanded to the final expansion stage.

We use a parameter d_{avg} to control the IP traffic load and create traffic matrices which – compared to $\mathcal{D}_{\text{fut}}$ – represent lower traffic requirements, e.g. taking place at earlier points in time. For the population-based traffic pattern, we set $d_{\text{tot}} = d_{\text{avg}} \cdot |\mathcal{V}| \cdot (|\mathcal{V}| - 1)$ and calculate d_{xy} according to Section 3.3. For the homogeneous traffic pattern, we set $d_{\text{hom}} = d_{\text{avg}}$. Hence, d_{tot} is identical for both traffic models, cf. Section 3.3.

The final number of IP demands that are to be routed in the IP topology for the four scenarios are illustrated in Fig. 5 for different values of d_{avg}. The figure shows that the population-based traffic pattern causes a slower increase of IP traffic than the homogeneous traffic pattern over parameter d_{avg}. While the traffic load is the same for all links with the homogeneous traffic pattern, the population-based traffic pattern reaches its maximum if at least one link is fully utilized. Hence, the traffic load created by both traffic patterns is maximal regarding the number of demands that can be carried by the underlying multi-layer network if blocking of demands

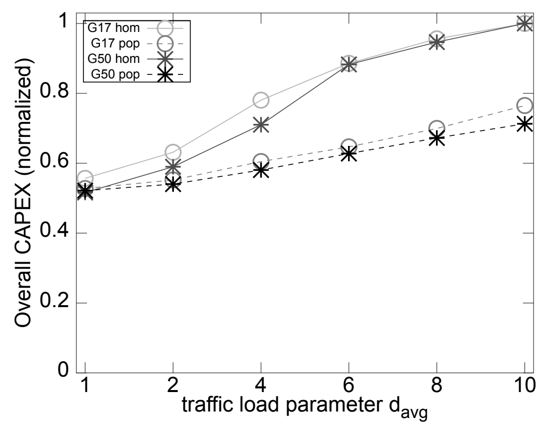


Figure 7: Evaluation of overall CAPEX for different traffic loads d_{avg}.

is not allowed.

Using the proposed algorithm, we performed several evaluations varying the traffic load by parameter d_{avg}. We considered different values for the parameter $w_{\max}$ in Equation (6) and found that $w_{\max}$ only has a minor impact in the considered scenarios. However, $w_{\max} = 1$ produced the best results on average and was chosen for the presented results. As the link candidate selection of Algorithm 1 given in Equation (5) is probabilistic, we perform $N = 10$ runs with different random seeds for each scenario. In each scenario, the best of the N runs is chosen. Fig. 6 shows the ratio of unused IP links to the number of links in the original IP layer topology, cf. Table 1. When the traffic load parameter d_{avg} is increased, the ratio of unused IP links decreases as expected. As we aforementioned, the maximal load for the population-based traffic pattern is reached before all links are fully utilized. Hence, the ratio does not drop to zero as it does with the homogeneous traffic pattern. In Fig. 7, the overall CAPEX is illustrated that can be achieved by reducing the topology according to our algorithm. Analogously to the unused IP links, the ratio of overall CAPEX increases as well when the traffic load parameter d_{avg} is increased.

Both figures show that the amount of edges as well as overall CAPEX significantly depend on the considered traffic pattern. Although the absolute number of demands routed per topology differs by far, the ratio of unused IP links as well as overall CAPEX only slightly differ per topology as shown in Fig. 6 and Fig. 7.

The only exception can be seen for G17 and G50 with homogeneous traffic pattern. The overall CAPEX value significantly differs for $d_{\text{avg}} < 6$ although the ratio of unused IP links is almost identical. This is explained by additional savings in the OMS and OCh layers. These savings occur in the G50 scenario since the amount of traffic between certain sites causes the installation of more powerful OMS equipment. The traffic requirements in the G17 network can be fulfilled with legacy OMS equipment. For $d_{\text{avg}} \geq 6$, the traffic load is sufficient to always require the more powerful OMS equipment in G50. This phenomenon cannot be seen for the population-based traffic scenarios since the traffic load distribution according to the sites' population causes the installation of powerful OMS equipment between sites with high population even for small values of d_{avg}.

6 Summary

In this paper, we presented an algorithm that is able to reduce a multi-layer network which was designed to carry future IP traffic requirements such that it fits to a given IP traffic matrix. The resulting network configurations enable incremental upgrades of the multi-layer network up to the final expansion stage.

We considered a homogeneous and a population-based traffic pattern on two German reference network topologies, designed corresponding multi-layer networks, and evaluated the proposed algorithm for these scenarios. The evaluations show that the ratio of IP links declared unused by our algorithm mainly depends on the considered traffic pattern and corresponds to the amount of saved CAPEX for the considered multi-layer network.

Acknowledgements

The authors would like to thank Prof. Tran-Gia for the stimulating environment which was a prerequisite for this work. Furthermore, we would like to acknowledge programming efforts by Eduard Weber and Julian Ott.

References

[1] Cisco Whitepaper, "Cisco Visual Networking Index – Forecast and Methodology, 2008–2013," Jun. 2009.

[2] M. S. Kodialam and T. V. Lakshman, "Dynamic Routing of Bandwidth Guaranteed Tunnels with Restoration," in *IEEE Infocom*, 2000, pp. 902–911.

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

[4] P. Pongpaibool and H. S. Kim, "Novel Algorithms for Dynamic Connection Provisioning with Guaranteed Service Level Agreements in IP-over-Optical Networks," in *IEEE Globecom* , 2003.

[5] R. Hüsermann, M. Jäger, S. O. Krumke, D. Poensgen, J. Rambau, and A. Tuchscherer, "Dynamic Routing Algorithms in Transparent Optical Networks ," in *Proceedings of the 7th IFIP Working Conference on Optical Network Design & Modelling (ONDM)*. Kluwer Academic Press, September 2003, pp. 293–312.

[6] M. Duelli, M. Hartmann, M. Menth, R. Hülsermann, and M. Düser, "Performance Evaluation of IP over Cost-Optimized Optical Multilayer Networks with SRLGs," in *9. ITG-Fachtagung Photonische Netze*, Leipzig, Germany, Apr. 2008, pp. 21–29.

[7] M. Duelli, E. Weber, and M. Menth, "A Generic Algorithm for CAPEX-Aware Multi-Layer Network Design," in *10. ITG-Fachtagung Photonische Netze*, Leipzig, Germany, May 2009.

[8] T. Michaelis, M. Duelli, M. Chamania, B. Lichtinger, F. Rambach, and S. Türk, "Network Planning, Control and Management Perspectives on Dynamic Networking," in *European Conference and Exhibition on Optical Communication (ECOC)*, Vienna, Austria, Sep. 2009.

[9] NOBEL Working Group, "NOBEL and NOBEL-2 Project Website," Sep. 2008. [Online]. Available: www.ist-nobel.org

[10] R. Hülsermann, M. Gunkel, C. Meusburger, and D. A. Schupke, "Cost Modeling and Evaluation of Capital Expenditures in Optical Multilayer Networks," *Journal of Optical Networking*, vol. 7, no. 9, pp. 814–833, Aug. 2008.

[11] M. Menth, "Efficient Admission Control and Routing for Resilient Communication Networks," Ph.D. dissertation, University of Würzburg, Jul. 2004.

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