Performance Evaluation of IP over Cost-Optimized Optical Multilayer Networks with SRLGs

Michael Duelli, Matthias Hartmann, and Michael Menth
University of Wuerzburg, Institute of Computer Science, Department of Distributed Systems
Am Hubland, 97074 Würzburg, Germany
E-mail: {duelli|hartmann|menth}@informatik.uni-wuerzburg.de

Ralf Hülsermann and Michael Düser
T-Systems Enterprise Services GmbH
Goslarer Ufer 35, 10589 Berlin, Germany
E-mail: {ralf.huelsermann|michael.dueser}@t-systems.com

Abstract

The design of optical transport and IP networks is investigated with respect to the effect on overall network performance when the two aspects are combined in the same network architecture. The focus is on the identification of advantages in the joint design of IP and transport networks, as well as on the shortcomings and network bottlenecks which may result from sub-optimal design. The impact of different IP routing strategies (shortest path, equal-cost multipath) on the network utilization is quantified, and potential sources for overload in the IP domain are determined. Simulations were carried out to systematically assess technical metrics like link utilization and economic parameters such as cost for a 50-node transport network with line rates up to 40 Gbit/s. The impact of fiber cuts was taken into account, leading to multiple logical IP links to fail simultaneously, i.e. shared risk link groups (SRLGs), and triggering massive IP rerouting. Applying the results allows to identify bottlenecks in the design, and to devise mechanisms which allow cost-optimal network design of future IP and optical transport networks.

1 Introduction

The introduction of new, mostly IP-based services in the past years has resulted in traffic volumes to continuously grow at rates of 50-100% per year. As this growth is anticipated to hold for the foreseeable future, cost-efficient physical transport technologies are required to handle Terabit/second network loads. Optical technologies such as the *optical transport network* (OTN) will provide the necessary high-bandwidth channels at rates of 40 Gigabit/second and beyond in future core networks.

In current IP networks, *multiprotocol label switching* (MPLS) and related protocols are used instead of native IP routing [1]. MPLS allows forwarding of packets based on labels which is faster than full IP address look-ups as well as the introduction of *traffic engineering* (TE) and *quality of service* (QoS) features. TE and QoS are essential to introduce load balancing, service differentiation, and resilience in traditional best-effort IP networks.

However, the cost of MPLS-based equipment is significantly higher than that of native IP and physical transport. This counteracts the carriers' need to decrease network costs. Hence, there has been renewed interest in using cost-efficient native IP protocols in combination with optical transport technologies, without the use of MPLS technologies in recent years. As this paper focuses on the potential of an interaction between pure IP and optical transport technologies, the currently widely used MPLS technologies were not considered.

We investigate two IP routing mechanisms, shortest path routing and *equal-cost multipath* (ECMP). These were applied to a 50-node transport network topology operating channels (links) at line rates of 2.5, 10, and 40 Gbit/s, and different levels of grooming. Higher line rates are more attractive since the relative capacity costs typically decrease with the increasing bandwidth of the interfaces. The connections established between terminating IP interfaces create the adjacencies between IP nodes and determine the topology of the IP layer. However, the potential cost savings of using transport technologies can be exploited only with high link utilization level, i.e. the IP layer routing directs a sufficiently high amount of traffic over the link to make this approach cost efficient. We demonstrate how the required capacity in the transport network can be optimized for a given traffic matrix, where the capacity could be in the granularity of OTH channels or even entire wavelengths.

Optimizing the use of different technologies in a future modular multilayer architecture is a complex problem: the network equipment and the routing on the network layer must be chosen such that the overall network costs are minimized. The challenge is in the trade-off of aggregating sufficient amounts of traffic on connections on the transport layer to achieve good resource utiliza-

The authors Michael Duelli, Matthias Hartmann, and Michael Menth have been funded by the Federal Ministry of Education and Research of the Federal Republic of Germany (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.

tion whilst avoiding overly long paths, which in turn would increase the transit traffic and overall traffic load in the network. We show the interrelation between the transport and IP network design process, both through the cost model as well as IP routing. The specific design dilemma addressed in this paper is as follows:

On the one hand, optical networks can be designed cost-efficiently, i.e. by knowing an edge-to-edge traffic matrix, connections on the transport layer of possibly different granularity are set up to carry the traffic between edge nodes. For this solution, it is a pre-requisite that traffic can be arbitrarily routed, in particular over multiple paths.

On the other hand, IP routing needs to obey constraints such as destination-based routing, load balancing limited to equal shares over equal-cost paths, and routing based on the shortest path principle which allows only for indirect TE through administrative link weights. In addition, IP reroutes traffic in failure cases such that overload may occur on backup paths.

Physical links often carry several transport connections. A fiber or cable cut may result in the failure of many IP adjacencies, which are referred to as *shared risk link group* (SRLG). IP reacts by rerouting traffic over backup paths in addition to the normal load on such paths. In case the necessary backup capacity was not carefully planned upfront, this often leads to congestion in the network. Therefore, we investigate cost-optimized network designs and study the effects of various IP routing variants, IP routing optimization methods, and potential failure sets. Our results show that cost-optimal network design must not ignore the constraints and assumptions of the routing layer since otherwise it is impossible to accommodate the traffic in the case of network failures.

The paper is structured as follows: Section 2 provides insight into the problem of transport network optimization, namely the question of grooming and cost-optimal design for a given topology and traffic matrix. The different IP routing mechanisms and IP network design principles are explained in Section 3 along with related work. This section also comprises an overview on IP optimization and native IP resilience. Section 4 provides data on the actual topologies and figures used for the quantitative analysis. The results of the different optimization exercises are presented and discussed under particular consideration of the problem of SRLGs. Conclusions from these results and an outlook on further work are given in Section 5.

2 Multilayer Network Modeling

2.1 Network Architecture

Incumbent network providers currently operate different network platforms for transporting IP traffic and for offering *synchronous digital hierarchy* (SDH) or *optical transport hierarchy* (OTH) based transport services. In most cases, IP and SDH/OTH platforms use a common fiber infrastructure but are operated as independent networks.

Traffic forwarding in the IP/MPLS layer is much more expensive than traffic switching on layer 2 or layer 1 [2, 3, 4]. Hence, it should be avoided that the whole IP traffic has to be switched in the IP layer at every node in the network when looking for cost-efficient network design principles for IP backbone networks. Typically, the IP routers are interconnected via high capacity *point-to-point* (p-t-p) links based on *wavelength division multiplexing* (WDM). Thus, the IP layer topology is determined by the configuration of connections in the optical layer. In the simplest case, the IP layer topology follows the physical network topology and all connections are terminated at each node in the network. We refer to this as *IP over p-t-p WDM* architecture.

With increasing end-to-end traffic demand between node pairs, a weakly meshed virtual IP layer network topology becomes increasingly less efficient since it implies a high amount of transit traffic in the IP routers leading to an unnecessary increase of the required IP router capacity. One idea for cost efficient IP backbone network design is to consider the traffic demand pattern even when designing the network topology and to install direct links between IP routers if the traffic demand between them is sufficiently high. In this case, the IP layer topology differs from the topology in the physical layer and traffic bypassing capability in lower network layers is needed for autonomous IP layer topology design.

There are a number of feasible multilayer network architectures providing traffic switching in different network layers and offering the opportunity for traffic bypassing. In [5], different architectures based either on transport network technologies (SDH/OTH), or optical switching, or a mix of both technologies are proposed. A systematic comparison of different architecture evolution scenarios assuming various traffic mixes focusing on network convergence is given in [6]. Results of quantitative case studies assuming realistic IP backbone network scenarios and different alternative multilayer network architectures are presented in [2] and [3]. An economic comparison of IP over WDM and IP over OTH architectures also addressing issues of multilayer restoration is given in [7].

Within the network modelling studies we present in this paper, an *IP over OTH* network architecture is assumed (Figure 1). In this scenario, the IP layer connectivity is realized via OTH connections and *electrical cross-connects* (EXC) and IP routers are installed in parallel. The EXCs are interconnected via point-to-point WDM links. The intermediate OTH layer provides several advantages - not all of those were exploited in the network architectures considered in this paper, though. First, traffic switching in the OTH layer is more cost-efficient

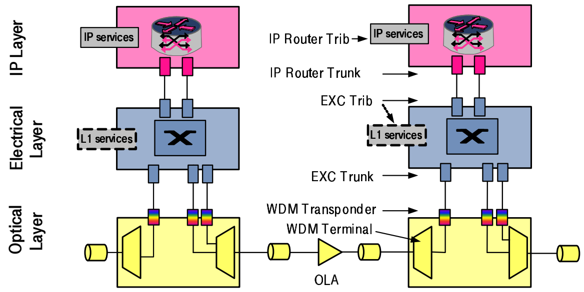


Figure 1: IP over OTH multilayer network architecture

and scalable compared to the IP layer. Second, flexible reconfiguration of OTH connections enables the operator to rapidly react to changing traffic patterns in the IP layer. Third, apart from savings in equipment investments, synergies from different services can be exploited: The use of the transport plane to also carry IP traffic will allow the operator to produce IP services and SDH/OTH transport services on a common platform. Furthermore, the incorporation of *automatic switched optical network* (ASON) or *generalized multiprotocol label switching* (GMPLS) based control plane functionalities could be used in real network architectures to implement such mechanisms.

2.2 Multilayer Network Design

Appropriate methods for multilayer network design and network dimensioning are needed in order to optimize network configuration and to minimize network costs. When assuming an IP over OTH network architecture, network design is related to the identification of the most cost efficient IP layer topology and therefore the configuration of OTH connections for a given network scenario. Hence, network design is targeted to the minimization of total network costs.

In the IP over p-t-p WDM reference scenario, IP routers are statically interconnected via WDM lightpaths without traffic bypassing capability, and the virtual IP layer topology follows the physical topology. Traffic demands are routed on the shortest path in this virtual topology and the capacity requirements are defined by the routing pattern.

In case of the IP over OTH network architecture, the virtual IP network topology is the result of an optimized traffic grooming procedure aiming at the minimization of the cumulated IP layer and electrical layer equipment costs.

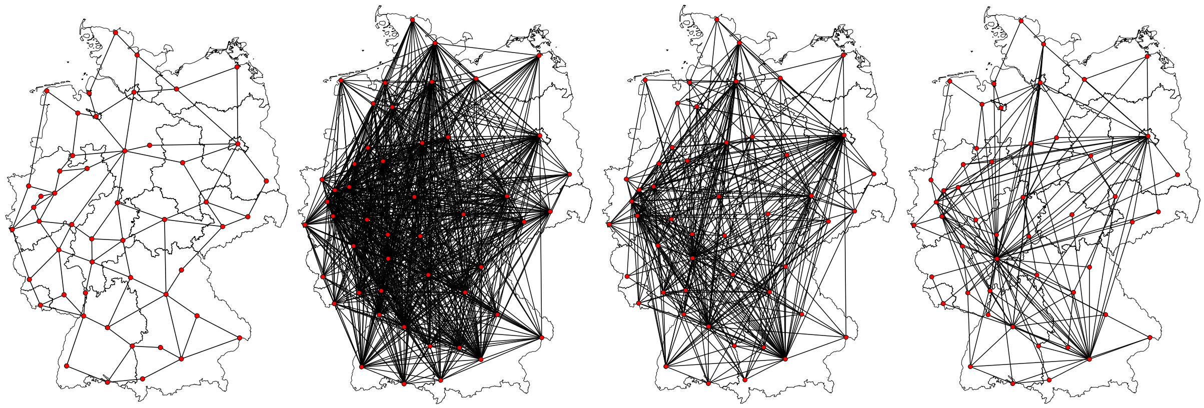
Traffic grooming refers to the process of grouping many small traffic flows into larger units. Efficient grooming is required in every multilayered network, regardless of the involved technologies. In [8], an overview of the traffic grooming problem and a survey on some representative work can be found.

Relating to the network design problem considered in this paper, grooming aims at the accumulation of end-to-end traffic flows in the IP layer in order to maximize the utilization of IP router interfaces and related OTH connections and to determine the most cost efficient IP layer topology according to the assumed traffic matrix. A variety of grooming algorithms related to different network architectures are described in literature [9, 10].

We use an end-to-end grooming algorithm [11] which is briefly described in the following: The algorithm starts with a fully meshed IP layer topology. The capacity of the IP layer links (i.e. granularity of OTH connections) equals the grooming line rate L and it can be set to ODU1 (2.5 Gbit/s), ODU2 (10 Gbit/s) or ODU3 (40 Gbit/s). If a particular traffic stream exceeds the connection bandwidth, IP link bundling, and parallel OTH connections are used. The algorithm tries to minimize the number of IP layer trunk interfaces by successively rerouting traffic flows which have been initially routed along weakly utilized IP links. In case of success, the free IP links and related OTH connections are released. It is allowed to split traffic flows, originally routed along the same link, to multiple routes. Although a constraint based routing per traffic flow like it can be provided by the MPLS technology would allow for such distributed routing patterns, pure IP routing, based on the shortest path principle, is not able to forward traffic, which is directed to the same node, along different routes. Apart from the optimization of the IP router interface utilization, the algorithm aims also at the minimization of the EXC port count by taking into account the port occupation of the IP layer links in the electrical layer. For resiliency reasons the grooming algorithm pays attention to the node connectivity: In minimum, two IP layer network links belonging to different SRLGs must be directed to a particular node in the network. Thus, neither a single link failure in the IP nor in the OTH or WDM layer will cause a fragmentation of the network.

After designing the virtual IP layer topology, the final capacity dimensioning is derived from the routing and grooming results. The number of IP router tributary interfaces is determined by dividing the amount of add/drop traffic by the assumed tributary line rate and rounding it up to the nearest integer. The line rate of IP router trunk interfaces and corresponding EXC tributary interfaces equals the grooming line rate and their number depends on the optimized virtual topology. The number of EXC trunk interfaces is determined by the routing of OTH connections. Due to simplification reasons the line rate of the EXC trunk interfaces is given by the grooming line rate and no additional multiplexing of OTH connections is assumed. Multiplexing of OTH connections can provide additional cost savings in some cases since cost efficient EXC ports and WDM transponders working at higher line speeds could be deployed for EXC interconnection. No protection issues are taken into account during network dimensioning. Therefore, all numbers refer to pure working capacity.

A comparative network dimensioning case study was



(a) German network topology, #nodes = 50, #links = 89

(b) L = 2.5 Gbit/s, #IP-Links = 841

(c) L = 10 Gbit/s, #IP-Links = 393

(d) L = 40 Gbit/s, #IP-Links = 160

Figure 2: Basic network topology and optimized IP layer topologies related to different grooming line rates

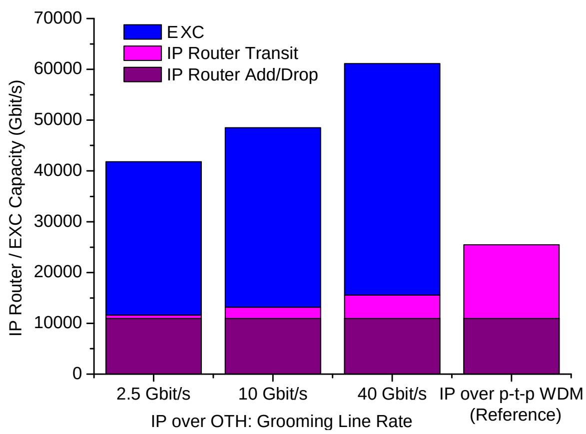


Figure 3: Overall net IP router and EXC capacity

performed based on a hypothetical but realistic German backbone network shown in Figure 2(a). End-to-end IP layer traffic demands are calculated on the basis of a population model taking into account statistical data on the regional population density and economic power [12, 13]. The total bi-directional traffic volume is about 2736 Gbit/s which relates to mean bi-directional traffic demand between node-pairs of about 2.23 Gbit/s. According to observations of the Internet traffic growth, this traffic volume is realistic in the near future [14]. The results for the total capacity of all IP routers and EXCs in the network are shown in Figure 3. In case of the IP router capacity, we have distinguished between the capacity needed to add and drop the traffic and the transit capacity. IP router add-drop capacity has to be equal in all scenarios. When comparing the IP/MPLS router capacity needed to handle transit traffic significant differences can be found: in the reference scenario more than 50% of the IP router capacity is dedicated to transit traffic. In case of the IP over OTH network architecture IP layer transit traffic is reduced significantly: with finer granularity of the OTH connections (i.e. the grooming line rate) the connectivity of the IP layer topology increases and the IP layer transit traffic decreases. If the grooming line rate is small, direct links between IP nodes are more attractive due to the higher interface utilization. It is obvious that the reduction of IP router capacity comes at the cost of additional EXC capacity. The difference in required EXC capacity among the different grooming line rates is caused by the decreasing utilization of OTH connections when going to higher grooming line rates. Figures 2(b)-2(d) show for each grooming line rate the related virtual IP layer topologies. As the grooming procedure aims at the introduction of direct links between nodes, the meshing degree of the virtual topologies is significantly high, depending on the grooming line rate, although not reaching a fully meshed IP layer, i.e. 1225 IP-Links.

2.3 Cost Comparison

Results of network dimensioning case studies presented in the previous section have shown that the IP over OTH network architecture decreases IP router capacity significantly but at the cost of additionally required OTH equipment. In order to quantify the economic benefit of the proposed architecture, the *capital expenditures* (CapEx) for installing the network infrastructure are evaluated.

The cost model is based on the model presented in [4] and is shown in Table 1. All cost values are normalized to the cost of a *long haul* (LH) 10 Gbit/s WDM transponder. The remarkable cost increase from the 640 Gbit/s to the 1280 Gbit/s IP router basic node is caused by the fact that all routers with capacity greater than 640 Gbit/s are multi-chassis configurations. While at the IP router tributary side 10 GE *(Gigabit Ethernet)* interfaces are assumed, PoS *(packet over SONET)* interfaces are used for IP router / EXC interconnection. As IP routers, EXCs, and WDM transponders are typically co-located, *short reach* (SR) interfaces can be used for interconnecting the devices. A WDM link is modeled by transponders, WDM terminals (i.e.

Table 1: Cost model used for CapEx analysis

Component	Costs
IP Router Basic Node 640 Gbit/s	16.67
IP Router Basic Node 1280 Gbit/s	111.67
IP Router Interface, PoS, 2.5 Gbit/s, SR	2.03
IP Router Interface, PoS, 10 Gbit/s, SR	6.88
IP Router Interface, PoS, 40 Gbit/s, SR	25.17
IP Router Interface, 10GE, Long Reach	3.34
EXC Basic Node 640 Gbit/s	13.33
EXC Basic Node 1280 Gbit/s	26.67
EXC Basic Node 2560 Gbit/s	69.33
EXC Interface, ODU1 (2.5 Gbit/s), SR	0.25
EXC Interface, ODU2 (10 Gbit/s), SR	0.67
EXC Interface, ODU3 (40 Gbit/s), SR	1.67
WDM Transponder, 2.5 Gbit/s, LH	0.33
WDM Transponder, 10 Gbit/s, LH	1.00
WDM Transponder, 40 Gbit/s, LH	3.75
WDM Terminal, LH, 40 channels	4.17
OLA, LH, 80km span length	1.92

mux/demux and amplifiers), and *optical line amplifiers* (OLA). The reach type of all WDM layer components is LH (i.e. 750 km) which is appropriate for all links in the German network.

The results of the CapEx analysis are shown in Figure 4. The total network costs are dominated by the IP router interfaces, as expected from the cost model. In case of the IP over OTH network architecture, the costs of IP router trunk interfaces can be significantly decreased and with this also the total network costs are reduced compared to the reference scenario although additional expenditures for EXCs are needed. Highest savings (≈ 28%) can be achieved when grooming the traffic into ODU2 connections (i.e. grooming line rate = 10 Gbit/s). If the grooming line rate is set to 2.5 Gbit/s (ODU1) the IP router costs are further reduced compared to the 10 Gbit/s scenario due to the lower transit traffic, but a huge number of 2.5 Gbit/s EXC interfaces are needed which comes at higher costs in total. Furthermore, the cost of WDM equipment is increased in the 2.5 Gbit/s scenario as the number of wavelength channels is highest. Additional multiplexing of ODU1 connections is recommended in order to deploy cost efficient EXC interfaces and WDM transponders working at higher bitrates and to decrease the number of wavelength channels in the 2.5 Gbit/s scenario. In case of the 40 Gbit/s scenario, the IP router costs are still very high. Much traffic grooming in the IP layer is needed in this scenario in order to efficiently utilize coarse granular IP layer links. Hence, a high amount of IP router trunk interface capacity is necessary which impacts the overall network costs.

A schematic representation of the interactions among multilayer network design parameters and outcomes is shown in Figure 5. The parameters considered in the

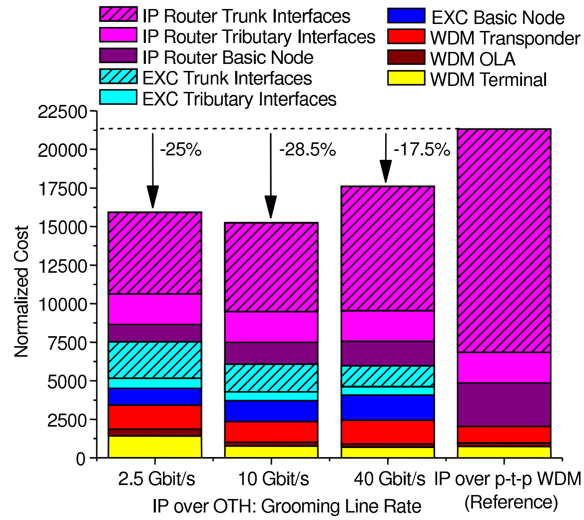


Figure 4: Results of CapEx analysis

planning framework so far are limited to the network scenario, network costs and a number of boundary conditions regarding network connectivity and the grooming linerate. The process is mainly targeted to the minimization of network costs. Nevertheless, within the network design phase the connectivity of the IP layer is determined which implies traffic routing on the IP layer. Furthermore, the routing of IP links in the physical topology defines SRLGs and therefore the IP layer performance in case of physical link failures (which show highest failure probability) gets also influenced by multilayer network design. These aspects are analysed and discussed in detail in the following sections.

3 Design of Optimal IP Routing in Optical Networks

In this section, we look at the design options of IP routing over optical networks. First, we give a short introduction to basic IP routing. Then, we explain a method to optimize the routing in the failure-free network and for specific failure scenarios. Furthermore, we explain the importance of considering the underlying optical network, when analyzing network failures.

3.1 Basics of IP Routing

In an IP network, each subscriber has a unique identifier – the *IP address*. Every IP packet sent over the network contains the IP address of its destination. Routers have routing tables that map such destination-addresses to outgoing network interfaces, and forward all incoming packets according to this mapping.

Routing tables are constructed in a distributed manner by routing protocols like *intermediate system to intermediate system* (IS-IS) [15] or *open shortest path first* (OSPF) [16].

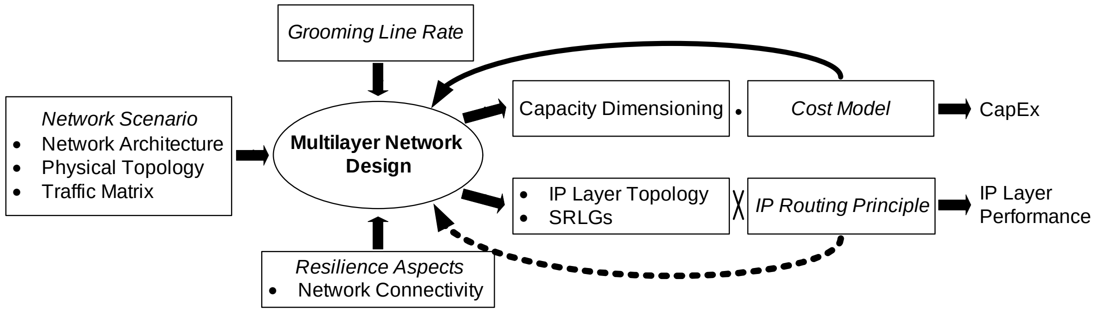


Figure 5: Combining the optimization of transport network design and IP routing

3.2 IP Routing with Multiple Shortest Paths

IP routing is intended to route traffic on the shortest path. When there are multiple shortest paths towards a destination, a router has to choose one of the corresponding outgoing network interfaces.

The most common approach in existing IP routers is to choose one of the shortest paths at random. As this means a loss of control and determinism in traffic engineering, this solution is accepted reluctantly at best by network providers.

There are several approaches to break this tie with multiple shortest paths in a deterministic way. In this paper, we investigate the two most important approaches. In shortest path routing (SPR), each network interface is associated with a unique integer identifier and the network interface with the smallest identifier is chosen in case of multiple shortest paths [15]. With *equal-cost multipath* (ECMP) [15, 16], traffic can be sent over all available shortest paths. We assume a perfect equal split of the data over all paths.

3.3 IP Reconvergence

The distributed calculation of shortest paths in IP networks is not static. Every time the network topology changes (through administrative changes or by failures), the new information is distributed to all routers. Based on this information, the routing is recalculated and converges to the new configuration. This *IP reconvergence* process can re-establish connectivity, as long as the network is not completely separated.

When analyzing failures, we only consider plain IP reconvergence. More sophisticated IP protection mechanisms based on *IP fast reroute* (IP-FRR) [17, 18] are not yet standardized and currently out of scope.

3.4 IP Routing Optimization

The shortest path algorithm which calculates the next-hops uses the administrative weights of each link as distance metric. These link weights are set by the network administrator. By modifying these link weights, we are partly able to engineer the traffic in an IP network. For example, highly loaded links can be made less attractive by increasing the weights of these links. To achieve an optimal routing of the given traffic demands through the network, several or even all link weights need to be adjusted.

In our optimization, we aim to minimize the maximum link utilization ρ of all links $l \in \mathcal{E}$. We analyze the capacity utilization $\rho(l, s)$ for a set of scenarios $s \in \mathcal{S}$, containing both the failure free scenario $\emptyset$ and certain failure scenarios.

$$\rho \quad = \quad \max_{s \in \mathcal{S}, l \in \mathcal{E}} \rho(l, s) \qquad (1)$$

Thus, ρ is the utilization of the link with the highest utilization in all considered failure scenarios.

A network can only handle the given traffic if ρ is below 1, otherwise packets are dropped at overloaded links and the network becomes unreliable. Thus, the routing should ensure that the load of the most congested link, and thus of all links, stays below 1. Therefore, we try to minimize ρ. Finding the minimum of ρ is an $\mathcal{NP}$-hard problem [19]. For networks beyond toy-size, heuristic methods are required to obtain at least near optimal results [20, 21, 22, 23], therefore. For our optimization, we use an implementation of the threshold accepting heuristic [24], as described in detail in [25].

3.5 Impact of Optical Networks

In an optical transport network, OTH connections are installed on top of the physical topology, cf. Section 2. They are typically routed over several physical links, and constitute a direct link between the start and end point in the IP layer.

The most likely failure in an optical network is the failure of a single physical link (e.g. by a fiber cut). A failed physical link causes the failure of all OTH connections that are routed over it. To analyze the performance of the virtual IP topology during failures, we must consider the combined failure of all these OTH connections. Thus, the set $\mathcal{S}$ contains one scenario s for each failure of a link in the physical topology. Each scenario $s \in \mathcal{S}$ is modeled as a failure of the related shared risk link group (SRLG), comprising all OTH connections that are mapped onto the failed physical link.

4 Optimizing IP Routing in Cost-Optimized Optical Networks

In this section, we merge the two worlds of designing cost-minimal optical networks and optimizing IP routing. We use the three IP layer topologies defined in Section 2 and compare them to the IP over p-t-p WDM scenario. We consider IP routing optimization in the failure-free case as well as all single link failures in the optical layer. Moreover, we consider failures in the pure IP layer.

For a comparison of our IP optimization, we need a reference configuration of the IP link weights. Therefore, we initially set all IP link weights to 1 which implies a usual *hop-count* routing, i.e. a routing with the smallest number of hops between source and destination. For this initial configuration, we calculate the maximum link utilization ρ according to Equation (1) for the topologies defined in Section 2. This value is listed in the first row of Table 2 grouped by topology for the routing variants SPR and ECMP.

In contrast to reality, values of $\rho > 1$ are also possible within our calculations. Therefore, the traffic matrix must be scaled by a factor $1/\rho$ to fit the traffic into the topology with the corresponding IP routing in reality.

4.1 Effect of IP Routing Optimization

To visualize the effect of IP routing optimization, we start with an illustration of the link utilization distribution in a network. We plotted the maximum link utilization for all links in the 40 Gbit/s topology with SPR in the SRLG case in Figure 6. We chose this scenario because it demonstrates the impact of optimization on the distribution of maximum link utilization very well.

In Figure 6, the maximum utilization of the links with hop-count routing, i.e. the reference value, is illustrated in descending order by the blue curve. We can see by this curve, that albeit almost all links are overloaded (> 1), a maximum link utilization of greater than 4 is only caused by relatively few links.

The green stems represent the maximum link utilization for all links after the optimization process. For comparison, the stems are plotted in the same order as before. We can see that not only the maximum link utilization has decreased from ≤ 10.38 to ≤ 3.48, but also that the load distribution of all links changed towards are more equally loaded network. For easier quantification, the distribution of the maximum link utilization after the optimization is illustrated by the red curve in an ordered manner.

4.2 Optimization Considering no Failures

During the construction of the considered topologies, only cost optimization with arbitrary routing was respected, but not IP routing. Thus, all values in Table 2 are significantly higher than 1, i.e. there is at least one link which is heavily overloaded. Consequently, a merely cost-optimized network is not well suited for IP routing. Both considered IP routing variants, SPR and ECMP, suffer from this lack. We can see that ECMP mostly yields better link utilizations than SPR. Thus, ECMP appears to be more more suitable for traffic engineering as it allows load balancing. The optimization results for the failure-free case for both variants is listed in the second row of Table 2.

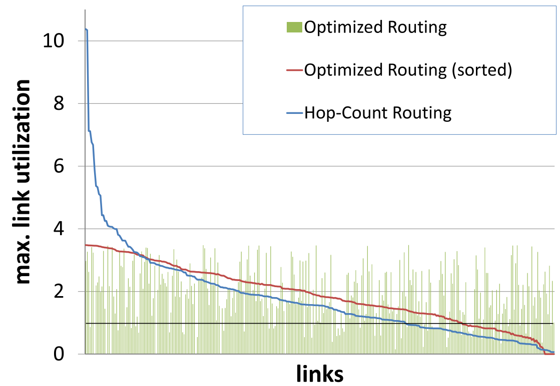


Figure 6: Maximum link utilization in the 40 Gbit/s topology in the SRLG case with SPR

We also optimized for the failure scenarios considering SRLGs. In the third row of Table 2, we have a look at the maximum link utilization changes in the failure-free case when optimizing for SRLGs. In general, we can see that our IP optimization is very effective. The maximum link utilization is often reduced by a multiple compared to hop-count routing. If we optimize the IP routing for the considered failures, the impact of the routing optimization degrades for non-failure cases as we can see in the third row of Table 2.

4.3 Optimization Considering SRLGs

Now, we also consider all single link failures in the optical layers, i.e. SRLGs occur in the IP layer as several IP links use the same optical link.

Table 3 incorporates the same structure as Table 2 for failure scenarios considering SRLGs. Although, twice the maximum link utilization could be anticipated for the failure cases since the minimum node degree in the physical topologies is equal or greater than 2, this factor is at about 5 to 10. The reason for this discrepancy is that the topologies were not designed for any failures at all and that a single physical failure can cause an SRLG affecting up to 100 IP links in the 2.5 Gbit/s topology.

If we optimize IP routing for handling these failure cases, the maximum link utilization for SRLGs improves. However, the optimization for failures bears the danger that the maximum link utilization increases in the failure-free case, which is evident in the third row of Table 2. This results from the lack of extra capacity

Table 2: Maximum link utilization ρ when considering failure-free scenarios

Architecture	IP over OTH						IP over p-t-p WDM (Reference)	
Grooming line rate	2.5 Gbit/s		10 Gbit/s		40 Gbit/s			
Routing Method	SPR	ECMP	SPR	ECMP	SPR	ECMP	SPR	ECMP
hop-count routing	7.33	1.78	6.43	1.84	2.96	1.84	6.71	10.98
optimized for failure-free	4.09	1.70	2.50	1.57	1.57	1.50	1.09	1.08
optimized for SRLG failures	11.19	7.56	5.49	4.22	2.69	2.49	6.71	7.22

Table 3: Maximum link utilization ρ when considering single physical link failures related to SRLGs

Architecture	IP over OTH						IP over p-t-p WDM (Reference)	
Grooming line rate	2.5 Gbit/s		10 Gbit/s		40 Gbit/s			
Routing Method	SPR	ECMP	SPR	ECMP	SPR	ECMP	SPR	ECMP
hop-count routing	42.54	15.94	14.08	8.99	10.38	4.97	37.38	33.86
optimized for failure-free	41.95	15.94	18.29	13.70	4.48	5.93	39.20	37.00
optimized for SRLG failures	15.94	15.94	7.18	6.37	3.48	3.38	8.69	8.69

in the cost-optimized OTH network design for this case. Thus, it is complex to make the right decision with pure IP routing in merely cost-optimized OTH networks.

4.4 Protecting IP Layer Failures

In optical networks, it is also possible to have optical protection by additional equipment which is able to almost seamlessly keep up the communication in case of a fiber cut. Amongst additional CapEx, these lack the ability to detect failures on higher layers, e.g. the IP layer. Thus, a failure of an IP port card is not protected. Hence, we only consider the mere resilience by the IP layer as it is able to handle both kinds of failures.

Opposed to a physical failure, a failure in the IP layer does not affect a whole SRLG. Just a single IP link fails, instead of several, and the increase of the maximum link utilization by this IP failure is rather small.

4.5 Summary

Recapitulating the results of the IP optimization on cost-optimized optical topologies, the results of Table 2 and Table 3 show that we succeeded in constructing an IP over OTH topology which outperforms the reference IP over p-t-p WDM topology w.r.t. CapEx as well as maximum link utilization, especially in case of failures. Additionally, we find that the information from the IP routing and IP routing optimization should be used to refine the multilayer design process in an iterative process as denoted in Figure 5 by the dotted feed-back arrow.

5 Conclusions

In this paper, we have discussed the design of cost-optimized IP over optical multilayer networks. The results are based on realistic assumptions for traffic matrix and equipment costs. We have shown that an IP over OTH network architecture and a cost-efficient multilayer network design provide significant cost savings compared to an IP over p-t-p WDM solution by moving transit traffic from IP towards the OTH layer.

The optimal network design assumes an unconstrained routing while IP routing needs to respect certain conditions. Therefore, we tried to answer the question: can we efficiently carry traffic over such a cost-optimal infrastructure using pure IP routing? Our results show that neither shortest path routing (SPR) nor equal-cost multipath (ECMP) routing can accommodate the traffic in the cost-optimized network: simple hop-count routing cannot fit the traffic into the pipes and also optimized IP routing leads to severe bottlenecks. Thus, when IP routing is intended to be used on the network layer instead of MPLS or Carrier Ethernet which enable explicit routing, then IP routing constraints should be respected in the optimization process.

The network was capacitated for failure-free scenarios without protection on the optical layer since restoration through rerouting on the IP layer was assumed. We considered cable cuts that lead to the simultaneous outage of multiple OTH channels such that multiple IP adjacencies fail. The worst case link utilization was a large multiple of the utilization in failure-free scenarios both with and without optimization of IP routing including the shared risk link groups (SRLGs). This shows that it is not possible by the routing algorithm to adequately distribute the load in failure cases and that a thorough backup capacity planning is required.

In future work, we intend to integrate IP routing constraints in the optimization of cost-optimal optical networks as well as backup capacity requirements due to rerouted traffic in case of cable or equipment failures.

References

[1] A. Farrel, *The Internet and Its Protocols: A Comparative Approach*. San Francisco, CA, USA: Morgan Kaufmann, 2004.

[2] R. Hülsermann, M. Jaeger, A. Betker, M. Düser, F.-J. Westphal, and A. Gladisch, "Quantitative Evaluation of Architectures for Multilayer Transport Networks," in *International Workshop on the Design of Reliable Communication Networks (DRCN)*, La Rochelle, France, Oct. 2007.

[3] S. Bodamer, D. Reviol, and J. Späth, "Resilient IP Backbone Network Solutions for a Realistic German Network Scenario," in *Proceedings of* 8^{th}*ITG-Fachtagung Photonische Netze*, Leipzig, Germany, May 2007.

[4] R. Hülsermann, M. Gunkel, C. Meusberger, and D. Schupke, "Cost Modeling and Evaluation of Capital Expenditures in Optical Multilayer Networks," *submitted for publication in Journal of Optical Networks*, 2008.

[5] S. Bodamer and J. Späth, "Modelling and Cost Analysis of IP/MPLS Backbone Networks Based on Optical Transport Technology," in *Proceedings of* 6^{th} *ITG-Fachtagung Photonische Netze*, Leipzig, Germany, May 2005.

[6] M. Jaeger, "Network Architectures for Future Optical Networks," in *International Conference on Transparent Optical Networks*, Jun. 2006.

[7] S. Sengupta, V. Kumar, and D. Saha, "Switched Optical Backbone for Cost-Effective Scalable Core IP Networks," *IEEE Communications Magazine*, vol. 41, no. 6, pp. 60–70, Jun. 2003.

[8] E. Modiano and P. Lin, "Traffic Grooming in WDM Networks," *IEEE Communications Magazine*, vol. 39, no. 7, pp. 124–129, Jul. 2001.

[9] J. Fang and A. Somani, "IP traffic grooming over WDM optical networks," in *Proceedings of IFIP Working Conference on Optical Network Design & Modelling (ONDM 2005)*, Feb. 2005, pp. 393–402.

[10] H. Zhu, H. Zang, K. Zhu, and B. Mukherjee, "A novel generic graph model for traffic grooming in heterogeneous WDM mesh networks," *IEEE/ACM Trans. Netw.*, vol. 11, no. 2, 2003.

[11] S. Verbrugge, D. Colle, M. Pickavet, and P. Demeester, "Capex comparison between link-by-link and end-to-end grooming in a European backbone network," in *11th International Telecommunications Network Strategy and Planning Symposium*, Jun. 2004, pp. 27–32.

[12] A. Dwivedi and R. E. Wagner, "Traffic Model for USA Long-Distance Optical Network," in *Proceedings of Optical Fiber Communication Conference (OFC)*, Baltimore, Maryland, USA, Mar. 2000, pp. 156–158.

[13] A. Betker and C. Gerlach, "Traffic Demand Model for National Optical Networks based on the Population Distribution," in *Proceedings of 4th ITG-Fachtagung Photonische Netze*, Leipzig, Germany, May 2003.

[14] A. M. Odlyzko, "Internet Traffic Growth: Sources and Implications," in *Optical Transmission Systems and Equipment for WDM Networking II*, B. B. Dingel, W. Weiershausen, A. K. Dutta, and K.-I. Sato, Eds., vol. 5247, Aug. 2003, pp. 1–15.

[15] D. Oran, "RFC1142: OSI IS-IS Intra-Domain Routing Protocol," Feb. 1990.

[16] J. Moy, "RFC2328: OSPF Version 2," Apr. 1998.

[17] M. Shand and S. Bryant, "IP Fast Reroute Framework," http://www.ietf.org/internet-drafts/draft-ietf-rtgwg-ipfrr-framework-08.txt, Feb. 2008, expires August 28, 2008.

[18] A. Atlas and A. Zinin, "Basic Specification for IP Fast-Reroute: Loop-Free Alternates," http://www.ietf.org/internet-drafts/draft-ietf-rtgwg-ipfrr-spec-base-11.txt, Feb. 2008, expires August 27, 2008.

[19] M. Pióro, Á. Szentesi, J. Harmatos, A. Jüttner, P. Gajowniczek, and S. Kozdrowski, "On Open Shortest Path First Related Network Optimisation Problems," *Performance Evaluation*, vol. 48, pp. 201 – 223, 2002.

[20] B. Fortz and M. Thorup, "Internet Traffic Engineering by Optimizing OSPF Weights," in *IEEE Infocom*, vol. 2, Tel-Aviv, Israel, 2000, pp. 519–528.

[21] B. Fortz, J. Rexford, and M. Thorup, "Traffic Engineering with Traditional IP Routing Protocols," *IEEE Communications Magazine*, vol. 40, no. 10, pp. 118–124, 2002.

[22] B. Fortz and M. Thorup, "Optimizing OSPF/IS-IS Weights in a Changing World," *IEEE Journal on Selected Areas in Communications*, vol. 20, pp. 756 – 767, May 2002.

[23] ——, "Robust Optimization of OSPF/IS-IS Weights," in *International Network Optimization Conference (INOC)*, Paris, France, Oct. 2003, pp. 225–230.

[24] G. Dueck and T. Scheuer, "Threshold Accepting; a General Purpose Optimization Algorithm," *Journal of Computational Physics*, vol. 90, pp. 161–175, 1990.

[25] M. Menth, M. Hartmann, and R. Martin, "Robust IP Link Costs for Multilayer Resilience," in *IFIP-TC6 Networking Conference (Networking)*, Atlanta, GA, USA, May 2007.