Enabling Efficient Multi-Layer Repair in Elastic Optical Networks by Gradually Superimposing SDN

Jeremias Blendin*, Daniel Herrmann*, Matthias Wichtlhuber*, Matthias Gunkel§, Felix Wissel§, and David Hausheer*
* Peer-to-Peer Systems Engineering Lab, Technische Universität Darmstadt, Germany
Email: {jblendin|daniel.herrmann|mwichtlh|hausheer}@ps.tu-darmstadt.de
§ Deutsche Telekom Technik, Darmstadt, Germany
Email: {GunkelM|Felix.Wissel}@telekom.de

Abstract—**Multi-layer resilience is one of the prominent new concepts for modern carrier networks; it efficiently combines the advantages of the optical and the packet layer. However, not all features offered by modern optical transport networks can be fully used by the packet layer yet. In case of a fiber cut, an optical protection mechanism restores the original IP topology after a short transient time. But such an optical restoration is expected to use a new, longer light-path, which in turn might affect the optical capacity of the link. Bit-rate flexible optical transceivers are able to utilize the remaining optical capacity efficiently by adapting the network link capacity accordingly. However, the packet layer is not able to cope with fluctuating link capacities; often the policy is to rather shut down the link completely instead of using the remaining capacity. Consequently, this paper proposes a Segment Routing-based approach with superimposed Software-defined Networking (SDN) to allow the IP network to benefit from these new features. The minimally invasive, gradual deployment of the system is investigated, while keeping other proven and resilient technologies and systems unmodified. Using the topology and traffic matrix of a large German carrier the feasibility of such a deployment is evaluated.**

I. INTRODUCTION

Recent advances in optical networking have led to an increasing number of carrier networks adopting dynamically reconfigurable equipment and becoming elastic optical networks (EONs). EONs dynamically adapt to failure situations such as fiber cuts by redirecting light paths in the optical domain. However, changing the optical path may lead to changes of its characteristics. The commonly longer restored light path leads to a smaller optical signal-to-noise ratio (OSNR) influencing the applicability of high capacity modulation schemes on this link [1]. Bit-rate flexible optical transceivers are able to adapt the link capacity to the available modulation scheme, with higher modulation schemes allowing higher link capacities.

Changing link capacities in the optical domain imposes challenges in current IP networks, which need to react dynamically. The industry standard for the packet layer in carrier grade networks is MPLS-based routing operated on top of the optical layer. In MPLS-based network, often Equal Cost Multi-Path (ECMP) routing is used to facilitate both redundancy and load balancing. Variable link capacities raise two problems when using MPLS and ECMP in the packet layer: first, the bandwidth limitation might be visible to the optical layer only, and second, if it is visible by the packet layer, ECMP does not allow different weights per link. Using current techniques, the only available options are either to reduce the load on all parallel ECMP links, or to disable the smaller, restored link entirely to avoid overload. Both options lead to an underutilization of available capacity in the optical layer, which can lead to packet loss on heavily loaded links.

Consequently, this paper proposes to enable source routing by introducing Segment Routing [2] to existing MPLS routers and superimposing SDN functionality at the edge of the MPLS network. The logically centralized SDN control plane is used for interfacing with the controller of the optical network to receive information about link bandwidth changes. When a link with reduced bandwidth is detected, the SDN controller takes over the control of affected traffic flows at the network edge and splits them in into smaller flows. Thereby, a fine granular traffic steering in the core of the MPLS network is possible to compensate for the impaired link. However, currently available SDN controller software is not as mature as available MPLS technology. Therefore, keeping the reliability, resilience, and trust in existing carrier network technology requires a gradual and minimally invasive introduction of SDN technology. This is achieved by introducing SDN control on selected network nodes only, with a focus on nodes that through which sufficient network traffic passes to allow for the compensation of capacity reductions. To the best knowledge of the authors, this approach is the first one that enables the packet layer to actually exploit the gains in link capacity efficiency introduced by flexible optical transponders [3], [4].

The contributions of this paper are:

- A description of requirements for IP-optical cooperation when repairing link failures with subsequent capacity reductions in EONs.
- A minimally-invasive, gradually deployable, superimposing SDN-based approach that complements MPLS for steering traffic off a network link during capacity reductions.
- An investigation into gradual deployment strategies to determine the smallest number of changes required when introducing the system to existing networks.

The remainder of this paper is organized as follows: Related work is discussed in Section II. The requirements for the packet layer are described in Section III followed by the system design in Section IV. The design is evaluated and discussed in Section V followed by a conclusion in Section VI.

978-3-901882-85-2

II. RELATED WORK

The resource optimization potential of bit-rate flexible optical transceivers in carrier networks was investigated by Amar et al. [5]. However, the work discusses the optical layer only; it is not described how the capacity gains can be exploited by the packet layer. There is a number of related works focusing on multi-layer traffic engineering and selective rerouting, which can be categorized in approaches related to Generalized Multi-Protocol Label Switching (GMPLS) and more recent approaches based on Software-Defined Networking (SDN) and are discussed in the same order.

GMPLS is an extension to the MPLS protocol. First multi-layer approaches were presented by Iovanna et al. [6] and Vigoureux at al. [7] with the objective of providing a unified control plane for packet switched and path switched networks. However, the approach of GMPLS led to a highly complex platform [8]–[10]. Moreover, GMPLS does not allow for a gradual deployment in carrier networks, which makes adoptions difficult in the conservative carrier community [11]. Despite GMPLS being developed for a long time and by many partners, it has not been widely adopted [10].

The Path Computation Element Communication Protocol (PCEP) [12] offers router-initiated path computation by a remote controller. However, reacting to capacity impaired links requires controller-initiated path updates as well as adding flow-splitting rules on the routers. Both features are not available in the current version of PCEP. The Application-based Network Operations Architecture (ABNO) [13] relies heavily on PCEP and proposes the ABNO controller that handles path computations on multiple network domains. Even though the design is similar in complexity to GMPLS, ABNO is currently discussed for standardization in the IETF [13]. Another approach using a unified control plane is presented by Liu et al. [8]. The system introduces an SDN controller and makes all network elements OpenFlow compatible by introducing an abstraction layer for non-OpenFlow hardware like optical switches. Azodolmolky et al. [14] investigate the problem from a testbed management perspective. The approach follows the augmented model by using the user-network interface (UNI) to communicate between the different control planes. A more in-depth discussion of software defined networking (SDN) based multi-layer approaches has been published by Shirazipour et al. [15].

Finally, MPLS RSVP-TE [16] could be used to implement selective rerouting of flows after a link failure as proposed e.g. by Gerstel et al. [17]. However, RSVP-TE cannot be easily combined with ECMP. A recent proposal in the IETF to introduce this feature [18] has expired with no successor. RSVP-TE has to be used for all traffic flows and paths in the network, which leads to a very high number of label paths compared to ECMP, where only one path is specified. Furthermore, supporting variable bandwidth links with RSVP-TE is still in the process of standardization [19].

In conclusion, GMPLS and existing SDN approaches introduce considerable complexity. Many recent SDN-based approaches required the entire network to be replaced by SDN-compatible technology. MPLS RSVP-TE is flow-based as the approach proposed in this paper, but the required features are not yet standardized. Consequently, we aim at an augmented SDN approach similar to the work by Azodolmolky et al. [14], as its complexity is considerably lower than the alternatives.

III. REQUIREMENTS FOR PACKET LAYER REPAIR OF LINK-CAPACITY REDUCTIONS IN EONS

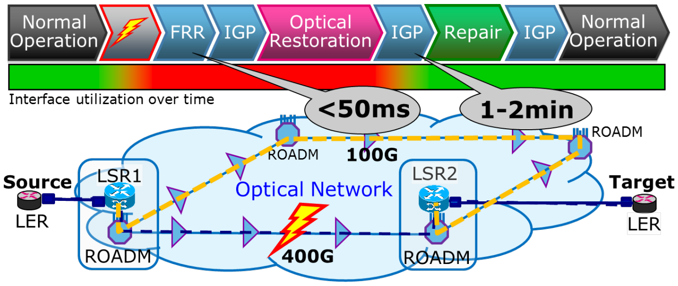


Fig. 1: An idealized view on an optical link failure and the subsequent restoration and repair (based on [20])

An example scenario for a link failure is depicted in Figure 1, which is also described in detail in [17]. The source Label Edge Router (LER) is connected to the target LER through the Label Switch Routers (LSR) 1 and 2. Both LSRs are connected to the optical network through a reconfigurable optical add-drop multiplexer (ROADM). Before the failure, the light path from LSR1 to LSR2 is using a direct fiber between the two sites and the optical link has a capacity of 400G. When the fiber breaks, the optical link fails, and triggers the fast reroute (FRR) mechanism that switches to an alternative MPLS path within 50ms (not shown in the figure). Then, the Interior Gateway Protocol (IGP) updates its routing tables and adjusts the traffic to the new conditions. Immediately after the broken fiber is detected, the optical controller starts the selection of a new optical path for the failed links. An alternative path is signaled and after a short delay, of typically between one and two minutes, it is physically established. The light path is much longer now and crosses two more ROADMs. Because of that the optical link offers a lower signal quality. Thus, the bit-rate flexible optical transceivers only provide 100G link capacities to the connected LSR1 and LSR2. At this point in time, a decision has to be made on how to use the restored optical link from the packet layer perspective.

Today, optical network recovery is a self-contained process of the optical layer. All major optical network vendors provide a logically centralized control plane. In general, these support both, an interface to query for information and a mechanism to send traps about changes to another controller. To be able to react on changes, the SDN controller needs to provide a southbound interface which can be connected to the control plane of the optical network. The optical control plane is assumed to provide information on the restoration process as well as the actual recovered link capacity. A closer integration between the optical and the packet layer seems not feasible, because of the prevalence of proprietary technology in optical networking. Link failures are understood to be events that occur regularly with the intervals being counted in weeks rather than days. Therefore, at most one fiber failure is expected to happen at the same time.

The requirements of carrier networks are demanding, especially in terms of resilience and costs. Resilience is particularly difficult to achieve in large distributed systems. Therefore, the resilience levels that are achieved today should not be reduced or impeded by the introduction of a new mechanism. The most widespread, highly resilient networking technologies that are available today in wide-area networks include MPLS. An alternative to relying on existing technologies like MPLS would be to switch the complete network to SDN. Although progress of SDN controllers in terms of resilience has been made, the area is still an active research topic [21] and no system is deemed to be up to the resilience levels of existing MPLS systems to this regard [22]. Furthermore, the lack of widespread deployments and therefore the lack of experiences and operational processes do not make SDN controllers the first choice for solely controlling production networks today. Therefore, approaches to increase the link utilization efficiency have to coexist with MPLS.

Generally, the deployment of new features in carrier networks is challenging. The existing operation of the system must not be affected and changes should be as small as possible. Therefore, it is not feasible to replace large parts of the network at once. Instead, an incremental deployment path must be considered. Consequently, the system has to be gradually deployable with minimal impact on the operation of the existing network [11].

IV. SYSTEM DESIGN

The core of carrier networks are often based on MPLS in the packet layer which is operated on top of elastic optical networks (EONs) to provide optical links as depicted in Figure 1. The superimposing SDN approach enables the packet layer to react efficiently to link capacity changes in the optical layer while satisfying the requirements described in Section III. MPLS relies on Label Switch Routers (LSR) in the core of the network. LSRs are optimized for high throughput and resilience and forward packets based on MPLS labels that have local significance only. In contrast, MPLS routers at the edge of the network, termed Label Edge Routers (LERs), determine the path of IP packets through the network. Consequently, LERs are the best candidates in a carrier network to gain control of the network traffic with the least amount of modification required in the network domain. The goal is to replace as few LERs as possible, while still gaining enough control over the network traffic to compensate link failures. The approach is described in more detail in [23].

A. *Concept*

A logically centralized optical controller interfaces with the SDN controller to provide link capacity changes in case of an optical restoration as depicted in Figure 2. To gain complete control over the path a packet takes through the MPLS network, the standard MPLS semantics of label switched paths are not sufficient. Therefore, Segment Routing (SR) [24], available as a software update for many MPLS enabled devices, is used. SR introduces new semantics for MPLS labels called segments that enables, in addition to other features, source routing. Packets are assigned segment stacks instead of a single MPLS label by the LERs. The relevant segment types for the superimposing SDN approach are node segments and adjacency segments. Node segments uniquely identify a node in the network as a packet's destination while adjacency segments address the link of a node and have local significance only. Once a segment has reached its destination, it is removed from the label stack and the next label is investigated for forwarding.

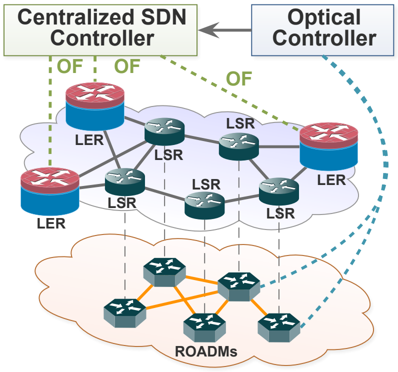
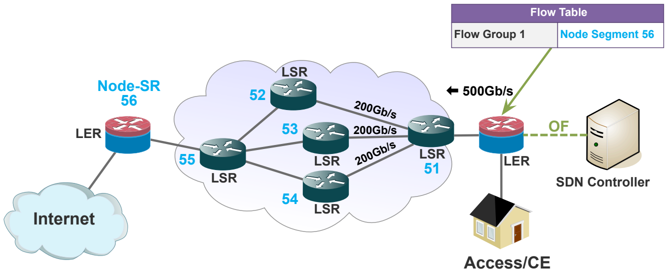


Fig. 2: Overview on the system design and its main components

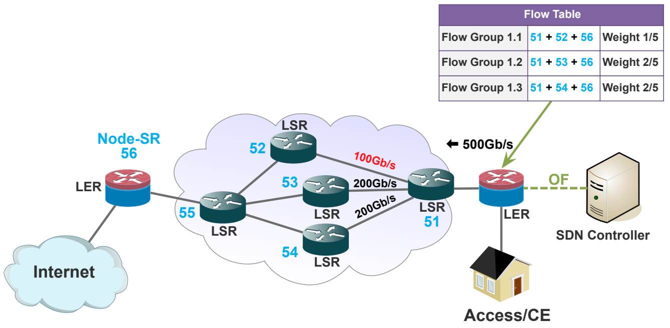
The normal operation of SR with node labels is depicted in Figure 3a and offers equivalent functionality to MPLS. Customer traffic bound for the Internet is received at an LER, which determines the egress LER towards the Internet. The corresponding LER has the node segment 56, which is added to each packet before forwarding it to the next hop, LSR with node segment 51 in the example. LSR 51 reads the topmost label of each packet's segment stack. In case of a node segment, the packet is forwarded to one of the next hops while the segment is left in place. The next hops for node 56 are provided by the underlying IGP and ECMP; the traffic is distributed evenly on the links to LSRs 52, 53, and 54. This process continues until the packets reach LER 56, where the node segment of the packets have reached their destination and are removed before they are forwarded out of the MPLS domain to the Internet uplink. The normal forwarding process does not involve the SDN controller at all; it uses the default, completely distributed SR mode of operation. While illustrated using SR, the default routing functionality is available through MPLS as well. When a link error occurs and the optical restoration results in a link with reduced capacity the SDN controller intervenes. Based on the state of the network the decision process comes to one of the following conclusions:

- (1) The reduced capacity on the link is sufficient to handle the traffic; no action is required.
- The reduced capacity is not sufficient to handle the traffic and the SDN control over the affected traffic is
  - (2) sufficient to compensate for the capacity reduction. The SDN controller superimposes the MPLS decisions and reroutes the traffic.
  - (3) not sufficient to compensate for the capacity reduction. The impaired link is disabled.

An example for a scenario with outcome (2) is depicted in Figure 3b. The capacity of the link between LSR 51 and LSR 52 is reduced from 200G to 100G, which is not sufficient for its ECMP load share of 125G. Using its knowledge of the state of the network, the superimposing SDN controller calculates a new load distribution for the affected flows in the network and



(a) High-Level overview on the system design including the main components



(b) High-Level overview on the system design including the main components

Fig. 3: Flow splitting and steering concept

applies it to the corresponding LERs. In the example, the SDN controller determines that the paths calculated by the IGP for ECMP are still usable, but their load shares need to be adapted. The OpenFlow protocol is used to override the decision of the MPLS and IGP process by adding an OpenFlow group table of type ”select”. This table enables ECMP-like traffic distribution but with weighted load shares per group table bucket on the LER where the affected traffic enters the network. Specifically, each bucket creates an individually steerable traffic flow. A different label stack is applied to each of the newly created flows that encode the three available paths from LSR 51 to LER 56. Thereby the load of the link with reduced capacity is reduced from 125G to 100G in the example, while the load on the other two links is increased to 200G. Using adjacency labels to encode the new paths is possible as well. However, using node labels retains the MPLS fast reroute feature [25]. Thereby, the approach retains the resilience and availability characteristics of the underlying MPLS network.

B. SDN Controller Architecture

The architecture of the SDN controller is depicted in Figure 4. The main component is the change controller that handles updates from the optical controller, runs the flow selection algorithm and decides which action has to be applied. The flow aggregation component processes the network state data and provides its results to the change controller. The relevant interfaces are described by their task in the remainder of this section.

In order to manage traffic flows and obtain the complete network topology, the SDN controller monitors the IGP messages exchanged by the routers. Either a network tap or a passive IGP node is used to collect data from the MPLS network. The data is used to derive a network state database that is identical to the one maintained by the routers in the network. Furthermore, the link utilization in the MPLS network is continuously monitored using existing monitoring mechanisms such as NetFlow/IPFIX. Because of the smoothing effect of statistical multiplexing, link utilization information is not required to be low latency. The decisions of the SDN controller are applied through the OpenFlow interface to the respective LERs.

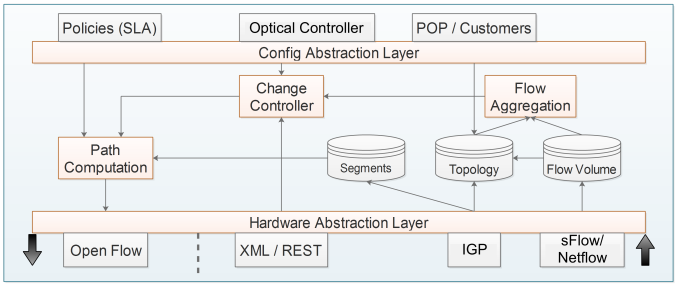


Fig. 4: High level overview on the Software Defined Networking controller

C. Interface between SDN Controller & LERs

In the system design and evaluation, the OpenFlow protocol is used for communications between the SDN controller and the SDN-enabled LERs. The main reason for that is the availability of OpenFlow in commercially available routers. Other protocols that support equivalent features can be used instead of OpenFlow to implement the approach; for example, PCEP [12] could be extended to include the required features. No matter what protocol is used, it is important that the SDN rules take precedence over the forwarding generated by the local MPLS process to ensure that they are applied only to relevant flows. The SDN controller selects flows for rerouting to compensate for the reduced link capacity. In this context, a flow is defined as the traffic that is forwarded from one LER to another LER. In carrier networks, the destination IP prefix of a packet usually determines the LER where it leaves the network. For each flow, on the source LER this process is implemented with OpenFlow by matching the destination IP prefix and forwarding the processing to the corresponding group table. For each flow a table of type ”select” is created on the originating LER. The number of buckets in the group table is determined by the number of links used by ECMP on the router adjacent to the reduced capacity link. Each bucket creates an independently steerable LER-to-LER flow. The steering is achieved by applying a segment stack that contains the list of all nodes segments on the selected path through the network.

D. Flow Selection

The flows to be rerouted are selected by an algorithm running in the change controller component of the SDN controller. As described in the preceding section, flows are steered unidirectional while fiber damages usually are bidirectional. Therefore, the algorithm models the network as a directed graph where two edges are affected by a link capacity reduction. For each edge in the directed graph, the algorithm described in this section is executed.

First, the topology database of the controller is updated to reflect the reduced capacity of the link. The bandwidth monitoring component of the SDN controller maintains utilization statistics for each link in the routing domain. Using this information and the link capacities, the algorithm can calculate whether an action is required or not using a simple comparison. If the measured bandwidth on the link before the failure is higher than the new capacity, action is needed; otherwise the optical controller is instructed to immediately activate the optical link. The network state information available to the SDN controller is used to determine the ECMP sub flows that pass the reduced capacity link. Furthermore, the system is designed to work with partial SDN control and thus not all LERs are controlled by the SDN controller. A list of flows that are controllable by SDN and use the reduced capacity link is generated. The algorithm selects the best flow for rerouting by sorting them by their bandwidth. The flows are used in sorted order until either the traffic on the link has been reduced to the required threshold or the list is empty. As depicted in Figure 5, in the latter case the algorithm aborts, because it is not possible to redirect enough flows. After selecting suitable flows, the algorithm performs sanity checks on them. Most importantly there must be a set of paths that does not use the reduced capacity link and are able to carry the additional traffic. In contrast to the calculation above, multiple paths do not need to be equal in IGP distance. As the flows can be split in smaller flows, any number of paths can be selected. Those paths are enforced by using SR node label stacks for every hop. Thus, multiple paths can be used in parallel without the need to be equal in IGP distance.

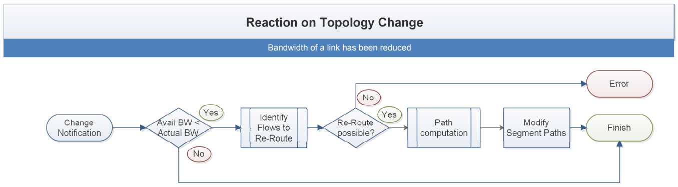


Fig. 5: Flowchart of redirection algorithm used to react on topology changes

V. EVALUATION

The goal of the evaluation is to twofold: first, to determine the minimum level of SDN-enabled devices in the core network required to control a sufficient amount of traffic to mitigate a link with reduced capacity. Second, the comparison of two deployment strategies those determine the order in which SDN features are deployed. To this end, the system is investigated using the topology of a large German carrier with its actual traffic matrix as well as a worst-case traffic matrix. The evaluation methodology is discussed in detail in Section V-A. The parameters used for the evaluation are discussed in Section V-B Finally, the results of the investigation of the gradual deployment is discussed in Section V-C.

A. *Methodology*

The evaluation investigates the interaction between a MPLS-based carrier network and an SDN controller that controls varying parts of the network's traffic. The system behavior in the event of a link capacity reduction is measured for a variable extend of SDN-control over LERs. Figure 6 shows a schematic overview of the network of a large carrier in Germany that is used as foundation for the evaluation. The core network consists of 12 points of presence (PoPs). A pair of core routers is located at each PoP that are connected to a self-redundant optical network. There are one or more aggregation rings per PoP, which connect LER locations to the core routers. Each LER aggregates the traffic of several broadband access gateways, which are finally connected to the customers.

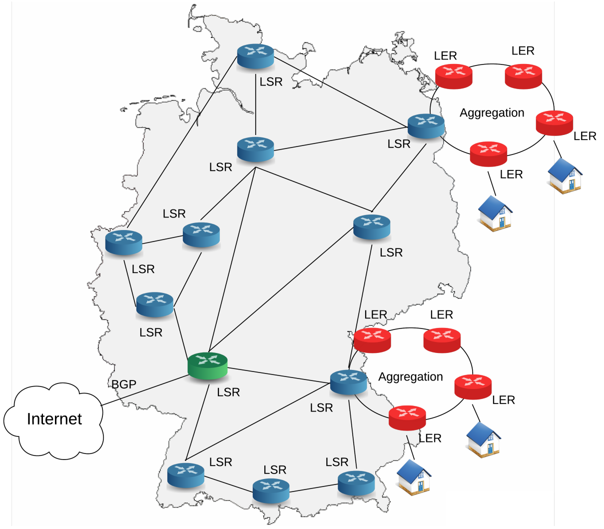


Fig. 6: The example network topology that closely resembles the core network of a large German carrier (based on [20])

For investigating the influence of SDN traffic steering at the LERs on the traffic between LSRs, only the links and PoPs are relevant. Each PoP is represented by an LSR that is complemented by a single, representative LER. In an actual network, the LSR is connected to an aggregation ring with a number of connected LERs. Different levels of SDN control over theses LERs is simulated by changing the ratio of SDN control on passing traffic on the representative LER. The resulting topology that is used in the evaluation consists of 12 core nodes (LSRs) with one edge node (LER) each. The level of SDN control in the core network is measured by the number of core nodes with attached edge nodes that can be partially or fully controlled by the SDN controller as well as the percentage of controlled traffic.

The investigated traffic matrices define how much traffic is sent between each pair LERs at PoP level. Their values are normalized to a maximum link capacity in the network of 100 Mbit/s to scramble the original values. The traffic generator assigns traffic randomly to IP addresses associated with each LER based on the traffic matrix. The flow splitting approach used for evaluation creates subnets for each IP range assigned to an LER. Each subnet, representing a 256th of the LER originated or addressed traffic, constitutes a flow that can be steered independently.

The link utilization both in default operations as well as in the case of faults is an important parameter. During default operations, the overall utilization must be low enough to allow every other link to fail without causing overload in the overall network. In the case of failure, links must increase their utilization to 100% to account for short-time traffic peaks. The approach used in this evaluation was introduced by Hasslinger et al. [26] and demands a maximum utilization of about 90%. Details on the formula can be found in the paper, in this evaluation the function is referred to as $\max\_\text{util}(x)$.

a) Trace-driven Analytical Model: The evaluation is conducted by calculating the link utilization in a network topology using a trace-driven analytical model. The model simulates the traffic distribution of ECMP and takes the non-uniform distribution of packets into account. Based on a PCAP file of traffic recorded from an Internet connection of a large research institution in Germany the analyzer calculates how the network traffic is split realistically by an ECMP algorithm. The ECMP algorithm calculates a hash value for each packet based on the IP source and destination addresses, the IP protocol number, and the layer 4 source and destination port, if present. The sum of these fields is calculated using per packet values from the PCAP file and is divided by the number of outgoing links, n. The remainder of this division $r \in [0, n-1]$ designates the link to transmit the packet on. For each number of parallel links (between 2 and 6) the summed packet length per interface is calculated. The resulting value is used to calculate the utilization of the links in the network when traffic is applied according to the traffic matrix.

B. Investigated Parameters

To determine the minimum level of SDN control and the influence of deployment strategies and other parameters is investigated. The used parameters, their values and value combinations are listed in Table I and described in detail in the following paragraphs.

TABLE I: Evaluated Parameters

Parameter	Values	Goal
Traffic matrix	ISP, worst case	–
Link failure scenario	most traffic, SDN control: highest, mean	–
% of reduced link capacity	50	–
Number of SDN-enabled PoPs	1,2,..12	Find minimum
SDN-control share of traffic per PoP	10,20,..,100	Find minimum
Deployment strategy	traffic, degree	Select better

a) Traffic matrix: Two traffic matrices are investigated, one carrier network traffic matrix and a worst case traffic matrix. For the carrier traffic matrix, the traffic distribution of the PoPs is derived from the actual traffic matrix of a large German carrier from 2015. This traffic matrix exhibits a Zipf-like distribution of traffic per PoP. The distribution is expected to be advantageous for a gradual deployment of the superimposing SDN approach. In contrast to that, the second traffic matrix represents the worst case for a gradual rollout, where the traffic is distributed evenly between PoPs.

b) Link failure scenario: This parameter describes the selection of the link that fails in this experiment. Three representative link impairment scenarios are investigated: the impairment of the link with the most total traffic, with the highest and with the mean controllability. The link with the most total traffic represents the worst case for the system as the highest traffic volume has to be rerouted to compensate the reduced link capacity. The link with the highest controllability represents the best case for the selected parameter set while the link with mean controllability represents the average of the other links in the network.

c) Percentage of removed link capacity: This parameter describes the capacity loss that is imposed on the capacity reduced link. While in general each value between 0% and 100% is possible, the values of 75%, 50% and 25% capacity reduction were investigated. Because of space constraints, this discussion focusses on 50% reduction.

d) Number of SDN-enabled PoPs: This number of SDN-enabled PoPs describes the threshold at which the system works reasonable well while still keeping the number of modifications to the existing network as low as possible. The network topology includes 12 core nodes, the SDN control is increased from 1 to the maximum number in steps of 1.

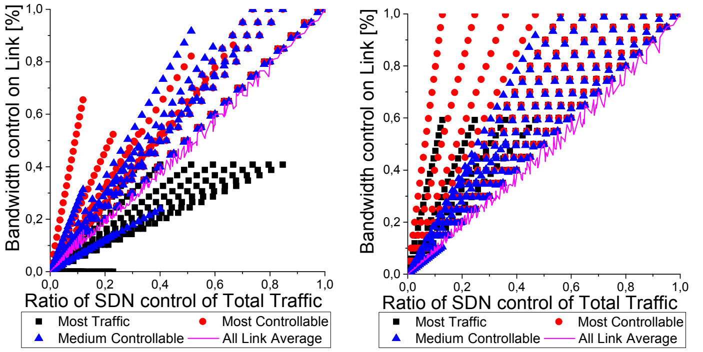
e) SDN-control share of traffic per PoP: Each core node in the evaluated topology represents a core PoP of the real core network. In each core PoP there are multiple access rings connected which are represented by a single LER per PoP in this evaluation. The number of SDN-enabled LERs in the ring is investigated by varying the amount of SDN control on the traffic per PoP between 0% and 100% in steps of 10%. A lower number of LERs is preferable due to fewer modifications that are required in the core network.

f) Gradual deployment strategy: Two methods on how to select the order in which the PoPs are SDN enabled are investigated. One approach is to deploy SDN to the PoPs in order of their originating traffic. More traffic at those nodes should result in more traffic in the overall network that can be controlled by the system. An alternative strategy is to deploy SDN to the PoPs in order of the number of their links. The number of links of a PoP, also termed the degree of the PoP, could indicate the influence of the node in the network. The first strategy is called traffic-based node selection; the second is called degree-based node selection.

C. Results

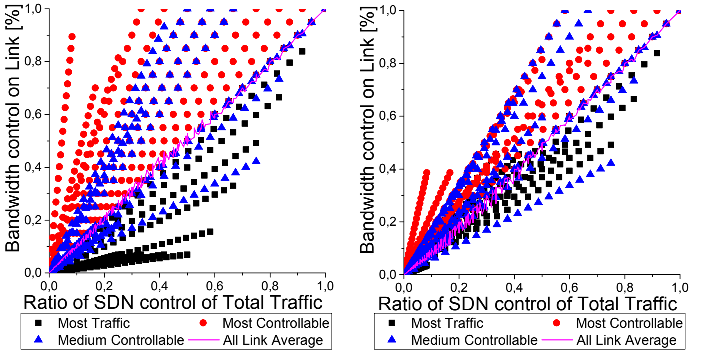
An aggregated comparison of on the SDN-control in the network and the share of controllable traffic on an impaired link is depicted in Figure 7. The traffic-based deployment strategy has clear advantages in the carrier traffic matrix depicted in Figure 7a. Because of the Zipf-like distribution of the traffic per PoP, the link with the most traffic can be controlled with only 50% SDN-control for the total traffic. The degree-based deployment strategy requires nearly 90% SDN-controlled traffic for the same result. In case of the worst-case traffic matrix this result is the other way round. Because of the fact that the differences in traffic per PoP are much smaller, the degree of a PoP becomes more important. The degree-based deployment strategy offers more efficient control over the link with the most traffic.

More detailed, per PoP results are depicted in Figure 8 for the carrier traffic matrix. Each graph shows the result of 360 experiments, 120 each for 3 different link failure scenarios, based on one deployment strategy. On the x-Axis, each graph shows the increasing number of SDN enabled PoPs (1 through



(a) Service provider traffic matrix, traffic-based node selection

(b) Service provider traffic matrix, degree-based node selection



(c) Worst case traffic matrix, traffic-based node selection

(d) Worst case traffic matrix, degree-based node selection

Fig. 7: Comparison of aggregated traffic control

12) and for each number also the percentage of SDN enabled traffic (10-100%). The order of the PoPs depends on the gradual deployment strategy, i.e. for 5 SDN nodes, the five largest PoPs either by traffic volume or by number of links are SDN enabled. For each experiment the gray bars in the background represent the percentage of the overall traffic in the core network that is SDN controlled. The right y-axis shows a normalized scale representing the global traffic in the network. A bar with a value of 1.0 on the right y-axis means that 100% of the total traffic is SDN controllable, which is only the case when all 12 nodes are 100% SDN enabled. The left y-Axis represents the absolute traffic on the impaired link. The three different colors represent the three different link failure scenarios. The solid lines represent the amount of traffic that needs to be removed from the particular link. This value is defined as $cap_{new} = \text{max\_util}(\text{cap} * \text{factor})$, given that cap is the original link capacity, factor is the removed capacity and $\text{max\_util}(x)$ is a function which calculates the usable link bandwidth as defined in Section V-A.

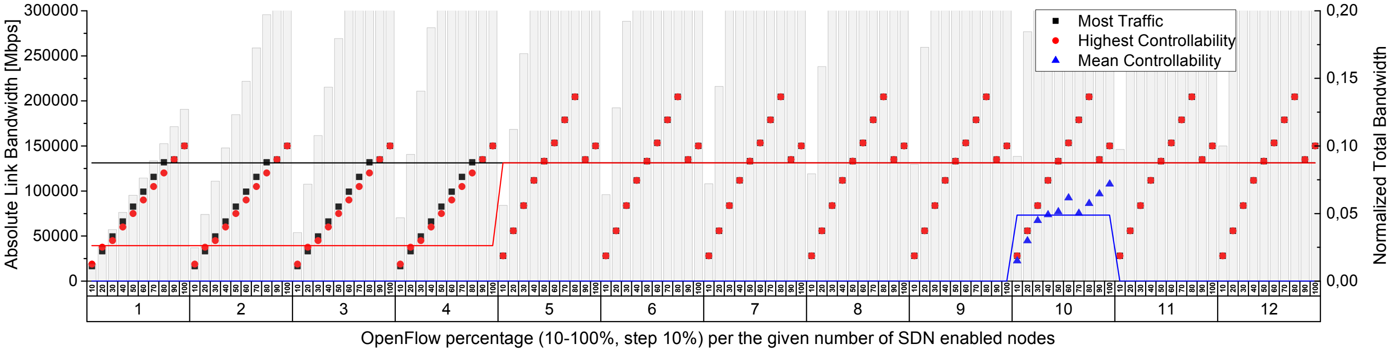
The outcome of an experiment, as described in Section IV-A, is indicated in the plots by the relation between the solid lines and the same colored dots. The amount of traffic that can be steered is plotted using the rectangular dots; they are plotted as absolute link traffic as indicated by the left y-axis. If a point is below the solid line, then the algorithm could not successfully redirect enough traffic to relieve the impaired link. Points above the solid line indicate a successful execution of the algorithm. The points do not represent the total redirect-able traffic, but the traffic that has actually been redirected. This means that there may be more SDN controlled flows, whose redirection is not required. For a given topology and traffic matrix, the link carrying the most traffic is the same in the experiments, as indicated by the identical black lines in all graphs. The red and the blue line represent the link that has the highest and the mean controllability through SDN at the given SDN configuration. This means that these links change with the number of PoPs that are SDN enabled. For example, the blue line represents the link with mean controllability. In Figure 7a it has as a non-zero value only at 10 SDN-controlled PoP nodes, in Figure 7b it appears with 5 and 6 SDN nodes and is zero otherwise. In other configuration the line has a value of 0, which means that no action is required to compensate for this link's impairment.

Figure 8 depicts results for the service provider traffic matrix and 50% capacity reduction of the failed link. The scenario depicted by Figure 8a shows the results for PoPs selected by the highest amount of traffic, while Figure 8b shows the degree-based selection. In Figure 8a, the steerability value for 5 SDN-enabled nodes at 90% of SDN-controlled traffic is lower than the value at 80%, which at first seems counter intuitive. This behavior reflects the fact that at this point enough traffic was redirected, and the link now is utilized below the threshold. Comparing both figures, the deployment order based on traffic per PoP has higher influence on the link with the most traffic. While the scenarios using 1 through 4 SDN PoPs only succeed in relieving the failed link when having 100% SDN controlled traffic, this changes when having 5 SDN-enabled PoPs or more. Already a percentage of 50% on 5 SDN nodes are sufficient to control the traffic with maximum utilization. The maximum controllable link equals the link with the highest utilization in this case, hence the lines and measurements are equal.

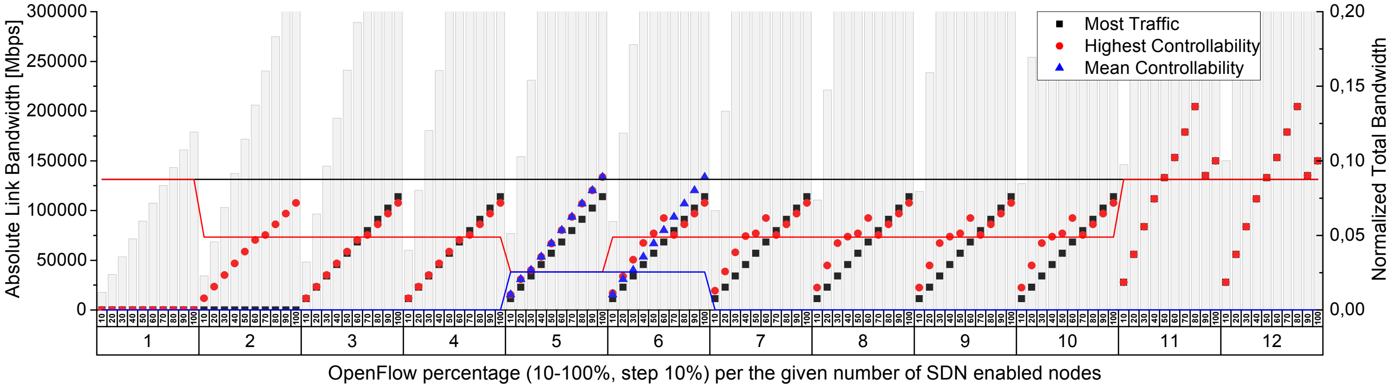
The Figure 8c show results for the worst case traffic matrix when using the traffic-based deployment strategy. Although the comparison of aggregated results in Figure 7 showed better results, the detailed view reveals that the traffic-based deployment yields better results than the alternative. The latter figure is omitted due to space constraints. The breakpoint for compensating both the most traffic and most controllable link impairment occurs with 7 SDN enabled PoPs. A slightly lower SDN control level per PoP is sufficient compared to the degree-based node selection method.

VI. CONCLUSION

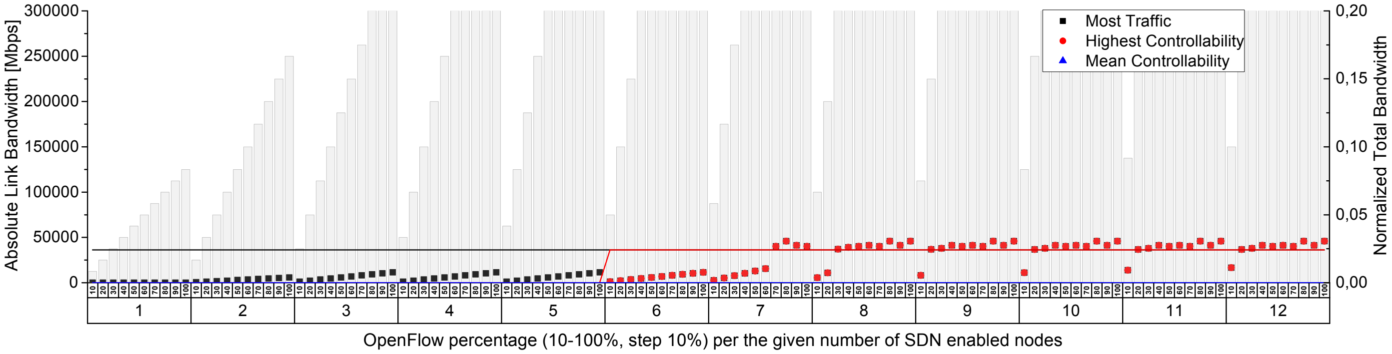
An approach to superimposing SDN to increase the efficiency of the packet layer in elastic optical carrier networks was introduced in this paper. Its ability to keep the resilience and reliability characteristics of proven MPLS technology is investigated while enabling centralized SDN control for failure recovery. If optical links offer reduced capacity after link failure, the SDN controller selectively overrides the MPLS infrastructure to utilize of the remaining link capacities more efficiently. Using the traffic matrix of a large German ISP, the system's ability to be gradually deployed while keeping most proven and resilient technologies and systems unchanged was shown. The most efficient approach to deploying the system



(a) Service provider traffic matrix, traffic-based node selection, and 50% link reduction on the failed link



(b) Service provider traffic matrix, degree-based node selection, and 50% link reduction on the failed link



(c) Worst case traffic matrix, traffic-based node selection, and 50% link reduction on the failed link

Fig. 8: Results of the investigation of gradual deployment strategies, the nodes on the x-axis are ordered by the respective deployment strategy

in a network with a Zipf-like per core PoP traffic distribution is to enable core PoPs ordered by their traffic volume.

It can be concluded that the packet layer has to be involved to exploit the potential of bit-rate flexible optical transceivers in elastic optical networks. The results show that gradually deploying new technologies to PoPs is a promising approach in the investigated network. Finally, superimposing SDN over an existing network is a viable approach to add new features to a network while minimizing its impact on the reliability characteristics of the network.

Next steps are the implementation of a prototype and its evaluation in a testbed. This includes the investigation of the costs of the system in terms of state in the network and computational requirements on the controller. Furthermore, the potential to reduce the required number of optical links, a major cost factor in carrier networks [20], should be investigated. Finally, the prototype should be investigated for its resilience characteristics with a focus on the interactions between the MPLS, the SDN and the optical controller.

ACKNOWLEDGMENT

This work has been supported by Deutsche Telekom (Dynamic Networks 4) and the German DFG (CRC 1053, MAKI). The authors would like to thank the reviewers, all their colleagues, and project partners for their valuable feedback and contributions.

REFERENCES

[1] M. Gunkel, F. Wissel, W. Weiershausen, M. Franzke, V. Fuerst, and A. Mattheus, “Multi-layer interworking with rate-adaptive transmission technology - benefit and challenges of a new use case,” in *ITG Symposion: Photonic Networks*, 2015.

[2] J. Tantsura, I. Milojevic, E. Crabbe, S. Previdi, M. Horneffer, A. Bashandy, R. Shakir, C. Filsfils, B. Decraene, and S. Ytti, “Segment routing architecture,” IETF, draft-ietf-spring-segment-routing-09, 2015.

[3] E. Riccardi, A. Pagano, E. Hugues-Salas, G. Zervas, D. Simeonidou, M. Bohn, A. Napoli, D. Rafique, A. D’Errico, N. Sambo, P. Castoldi, T. Rahman, M. Svaluto Moreolo, J. M. Fabrega, and M. Gunkel, “Sliceable bandwidth variable transponder: The IDEALIST vision,” in *European Conference on Networks and Communications (EuCNC)*, 2015.

[4] A. Lord, P. Wright, and A. Mitra, “Core networks in the flexgrid era,” *Journal of Lightwave Technology*, vol. 33, pp. 1126–1135, 2015.

[5] D. Amar, E. L. Rouzic, N. Brochier, and C. Lepers, “Multilayer Restoration in Elastic Optical Networks,” in *IFIP International Conference on Optical Network Design and Modeling (ONDM)*, 2015.

[6] 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.

[7] M. Vigoureux, B. Berde, L. Andersson, T. Cinkler, L. Levrau, M. Ondata, D. Colle, J. Fernandez-Palacios, and M. Jager, “Multilayer Traffic Engineering for GMPLS-Enabled Networks,” *IEEE Communications Magazine*, vol. 43, no. 7, pp. 44–50, 2005.

[8] L. Liu, D. Zhang, T. Tsuritani, R. Vilalta, R. Casellas, L. Hong, I. Morita, H. Guo, J. Wu, R. Martinez, and R. Muñoz, “First Field Trial of an OpenFlow-based Unified Control Plane for Multi-layer Multi-granularity Optical Networks,” in *Optical Fiber Communication Conference (OFC)*, 2012.

[9] A. Farrel, “A unified control plane: Dream or pipedream,” in *International Conference on IP+ Optical Network (IPOP)*, 2010.

[10] S. Das, G. Parulkar, and N. McKeown, “Why OpenFlow/SDN can Succeed where GMPLS Failed,” in *European Conference and Exhibition on Optical Communication (ECOC)*, 2012.

[11] D. Levin, M. Canini, S. Schmid, F. Schaffert, and A. Feldmann, “Panopticon: Reaping the benefits of incremental sdn deployment in enterprise networks,” in *USENIX Annual Technical Conference (USENIX ATC)*, 2014.

[12] J. Vasseur and J. Le Roux, “Path Computation Element (PCE) Communication Protocol (PCEP),” IETF, RFC 5440, 2009.

[13] A. Aguado, V. Lopez, J. Marhuenda, O. Gonzalez de dios, and J.-P. Fernández-Palacios, “ABNO: A Feasible SDN Approach for Multi-Vendor IP and Optical Networks,” in *Optical Fiber Communication Conference (OFC)*, 2014.

[14] S. Azodolmolky, R. Nejabati, E. Escalona, R. Jayakumar, N. Efstathiou, and D. Simeonidou, “Integrated OpenFlow-GMPLS Control Plane: an Overlay Model for Software Defined Packet over Optical Networks.” *Optics express*, vol. 19, no. 26, pp. B421–B428, 2011.

[15] M. Shirazipour, Y. Zhang, N. Beheshti, G. Lefebvre, and M. Tatipamula, “OpenFlow and Multi-layer Extensions: Overview and Next Steps,” in *European Workshop on Software Defined Networking (EWSDN)*, 2012.

[16] D. Awduche and Y. Rekhter, “Multiprotocol Lambda Switching: Combining MPLS Traffic Engineering Control with Optical Crossconnects,” *IEEE Communications Magazine*, vol. 39, no. 3, pp. 111–116, 2001.

[17] O. Gerstel, C. Filsfils, and W. Wakim, “IP-Optical Interaction during Traffic Restoration,” in *Optical Fiber Communication Conference/National Fiber Optic Engineers Conference*, 2013.

[18] K. Kompella, M. Hellers, and R. Singh, “Multi-path Label Switched Paths Signaled Using RSVP-TE,” IETF, draft-kompella-mpls-rsvp-ecmp-06.txt, expired, 2015.

[19] H. Long, M. Ye, G. Mirsky, A. D’Alessandro, and H. Shah, “Ethernet Traffic Parameters with Availability Information,” IETF, draft-ietf-ccamp-rsvp-te-bandwidth-availability-04, 2016.

[20] M. Gunkel and M. Horneffer, “Multi-layer Resilience in Deutsche Telekom’s IP Core and Aggregation Network,” in *ITG Symposion: Photonic Networks*, 2014.

[21] A. S. Da Silva, P. Smith, A. Mauthe, and A. Schaeffer-Filho, “Resilience support in software-defined networking: A survey,” *Computer Networks*, vol. 92, pp. 189–207, 2015.

[22] S. Scott-hayward, “Design and deployment of secure , robust , and resilient SDN Controllers,” in *IEEE Conference on Network Softwarization (NetSoft)*, 2015.

[23] D. Herrmann, “Underlay-Aware Traffic Engineering based on SDN,” Masters’ thesis, Technische Universität Darmstadt, 2015.

[24] C. Filsfils, N. K. Nainar, C. Pignataro, J. C. Cardona, and P. Francois, “The Segment Routing Architecture,” in *IEEE Global Communications Conference (GLOBECOM)*, 2015.

[25] P. Francois, C. Filsfils, A. Bashandy, B. Decraene, and S. Litkowski, “Service Function Chaining Problem Statement,” IETF, draft-francois-rtgwg-segment-routing-ti-lfa-00.txt, 2015.

[26] G. Haßlinger, J. Mende, R. Geib, T. Beckhaus, and F. Hartleb, “Measurement and characteristics of aggregated traffic in broadband access networks,” in *Managing Traffic Performance in Converged Networks*, 2007.