Greedy Design of Resilient Multi-Layer Networks

Michael Duelli, Xiaohua Qin, and Michael Menth
University of Würzburg, Institute of Computer Science, Chair of Communication Networks
Am Hubland, 97074 Würzburg, Germany
E-mail: {duelli,qin,menth}@informatik.uni-wuerzburg.de

***Abstract*—In this paper, we propose a deterministic greedy heuristic providing a construction layout for a cost-efficient multi-layer network that is able to carry a given set of traffic demands with and without protection on different layers. We apply the heuristic to different reference network topologies and protection requirements. Evaluations are conducted regarding equipment cost on different layers, blocking probability, path lengths, and number of demands affected by specific failures.**

I. Introduction

Network providers have been able to keep up with the ongoing exponential IP traffic growth [1] by means of equipment upgrades and the deployment of transport network technologies like *optical transport networks* (OTN) [2]. Each transport technology incorporates switching equipment and can be seen as a separate *layer* including its own topology within a *multi-layer network*. The use of a hierarchy of transport layers implies *shared risk groups* (SRG), e.g. all links using a common resource form a group sharing the risk of this resource's failure. SRGs increase the likelihood for contract penalty fees due to violation of *service level agreements* (SLA) and, hence, intensify the importance of resilience mechanisms in today's multi-layer networks.

In multi-layer networks, the installation of equipment enables logical links in higher layers to access physical link resources at specific bandwidth granularities as well as aggregation for improved resource efficiency. Resilience mechanisms require additional resources which already need to be addressed during network design. Nevertheless, resilient multi-layer networks have to be cost-efficient to ensure economic competitiveness. Since the cost-optimal design of resilient multi-layer networks is an $\mathcal{NP}$-hard combinatorial problem [3], mathematical approaches like *integer linear programs* (ILP) are not suited for comprehensive parameter studies on large network instances due to long computation times.

In this paper, we present a deterministic greedy heuristic for *capital expenditure* (CAPEX) aware design of resilient multi-layer networks. A physical topology, a network equipment and CAPEX model, as well as a set of traffic demands are required as input for the multi-layer network design. The result of the heuristic are construction layouts for all considered networking layers as well as multi-layer paths for as many of the given traffic demands as possible. We use the multi-layer network equipment and CAPEX model of the IST Nobel project [4] and consider dedicated path and link protection mechanisms on different layers for different multi-layer network instances and failure scenarios. We perform evaluations regarding the equipment cost on different layers, blocking probability of traffic demands, path lengths, and number of traffic demands affected by specific failures.

The remainder of this paper is structured as follows. Section II explains the models for multi-layer network equipment and CAPEX as well as resilience. In Section III, we outline the problem formulation and give an overview of related work. Section IV presents the heuristic for CAPEX-aware resilient multi-layer network design which is evaluated in Section V. Finally, we summarize this paper in Section VI.

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.

II. Multi-Layer Network and Resilience Modeling

This section describes the models for multi-layer network equipment, CAPEX, and resilience.

A. Network Equipment and CAPEX Model

A network equipment model defines the available technologies and their possible interconnections. In turn, a CAPEX model $\mathcal{C}$ associates cost values with each networking equipment. In [4], a network equipment and CAPEX model was given covering IP/MPLS, Ethernet, SDH/SONET, and *optical transport network* (OTN) technologies. We denote technologies as *layers* and define the set of layers $\mathcal{L}$ comprised in $\mathcal{C}$. The cost values are vendor-independent and normalized to the cost of a 10 Gbit/s WDM transponder. The authors of [5] list the available network equipment and corresponding CAPEX values for this model.

We split the equipment of all technologies considered in [5] into four modular component groups illustrated in Fig. 1. *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.

As described in [6], we strictly apply the four component model on all considered technologies in contrast to the original model given in [5]. Consequently, we provide zero-cost dummy components in case component groups were not defined in the original equipment model. Furthermore, we split

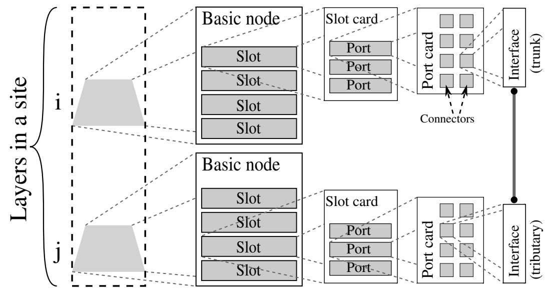


Fig. 1. Interconnection of two layers i and j within a site using modular components of the four component groups.

the OTN layer into an *optical channel* (OCh) layer dealing with separate wavelengths and lightpaths as well as an *optical multiplex section* (OMS) layer handling and switching bundles of wavelengths on the fiber strands.

B. Layer Interconnection Model

Within a layer, the assembly of the four modular component groups at a site 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 layers within a site or across remote sites requires that the transmitting interface encapsulates the data in a *compatible* format that can be processed by the receiving interface.

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 layer and ability, interfaces use a specific modulation and framing to en/decapsulate data. The framed data is transmitted with an interface specific bit rate, the interface's *capacity*. Two interfaces must use the same data encoding, i.e. modulation, framing, and capacity, to be compatible. With the interconnection of two layers, data might also be aggregated or deaggregated either by an interface, e.g. a muxponder, or during the switching process at the basic node, to increase resource utilization.

C. Multi-Layer Path Model

An interconnection of two sites across multiple layers and intermediate sites may involve repeated encapsulation and aggregation of data at each layer as well as the decapsulation and deaggregation of data at intermediate sites. Fig. 2 illustrates a *multi-layer path* between three sites on which data is en- and decapsulated several times such that data can be processed at the responsible layers.

Thus, the installation of equipment corresponds to the establishment of higher layer logical links in this model. These logical links are realized by underlying logical or physical links, in turn. As a consequence, the multi-layer path of a traffic demand is a concatenation of edges in logical layers which themselves are recursively defined by concatenations of logical or physical edges within the set of available layers $\mathcal{L}$. The recursive definition stops at the physical layer which interconnects the remote sites by physical fiber strands. An exemplary multi-layer path across three layers along four sites is depicted in Fig. 3. Therein, a logical connection of site A to D at layer i is recursively defined by two logical connections from site A to C and C to D defined in layer j and so on.

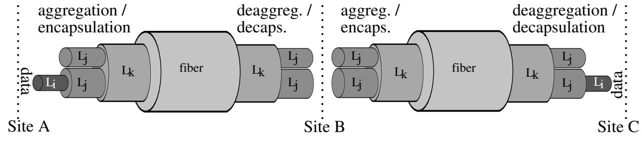


Fig. 2. Setup of a multi-layer path via data en/decapsulation across multiple layers and sites.

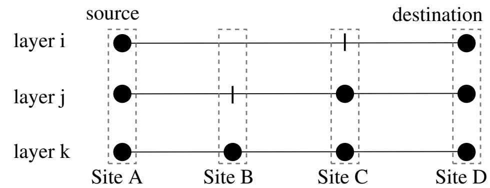


Fig. 3. Profile of an exemplary multi-layer path with four sites.

D. Resilience Model

Every component in the network equipment model can fail. As equipment realizes logical edges in our model, component failures can be mapped to corresponding link failures. Hence, we define the set $\mathcal{S}$ of all possible failures which is the powerset of all edges $\mathcal{E}_l$ in all layers $l \in \mathcal{L}$ of a multi-layer network. A special case is the failure of a whole site which affects all incident edges of this site on all layers.

In this paper, we focus on *dedicated protection* mechanisms in a single layer chosen of all layers $\mathcal{L}$. We consider *link* as well as *path protection*, i.e. setting up a backup path for each logical edge for local repair or a single end-to-end backup path, respectively. For path protection, we distinguish link and node disjointness of primary and backup paths.

III. Problem Formulation and Related Work

A formulation of the optimization problem as well as an overview of related work are given in this section.

A. Problem Formulation

For a given set of traffic demands $\mathcal{D}$, we design a resilient multi-layer network from scratch. Initially, merely the physical layer is given and the logical layers are empty, as depicted in Fig. 4. The physical layer is unalterable due to high cost of earthwork whereas logical layers can be arbitrarily modified. Links in logical layers originate from equipment that is installed at the sites. Thus, the multi-layer network is configured by deploying and interconnecting networking equipment to provide resources for primary and backup paths of traffic demands. We merely know that the source and destination of a traffic demand is to be connected at its originating layer whereas the traversed sites and required resources have to be determined. Additionally, the installation of further equipment allows us to not only make use of existing edges in logical layers but also create new edges between any pairs of sites. The CAPEX-aware deployment of equipment

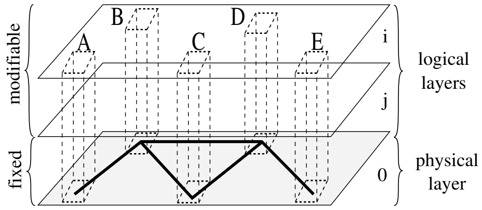


Fig. 4. AN EXEMPLARY MULTI-LAYER NETWORK CONSISTING OF A GIVEN PHYSICAL TOPOLOGY AND INITIALLY EMPTY LOGICAL LAYERS.

to fulfill the requirements of the traffic demands is an $\mathcal{NP}$-hard combinatorial problem [3] whose complexity basically depends on the number $|\mathcal{L}|$ of layers/technologies, the number $|\mathcal{V}|$ of sites/*points of presence* (PoP), and the number $|\mathcal{D}|$ of traffic demands. This complexity can be roughly estimated by

$$\mathcal{O}\left(2^{|\mathcal{L}|} \cdot |\mathcal{V}| \cdot |\mathcal{D}|!\right) \quad (1)$$

as $|\mathcal{L}|$ technologies may be either used or not used at any of $|\mathcal{V}|$ sites per demand while the optimal order of traffic demands can be found within $|\mathcal{D}|!$ possible permutations.

For resiliency, two disjoint multi-layer paths are required for each demand. The constraint regarding disjointness of primary and backup paths for the considered resilience requirements must hold in all lower layers, including the physical layer. The resilience requirement further increases the complexity of the considered problem.

In principle, this problem can be solved by mathematical approaches, but their usage is limited in practice due to long computation times. In this paper, we propose a deterministic and fast heuristic that is also flexible regarding the network equipment model. The heuristic does not approach the combinatorial problem but creates edges in logical layers which imply equipment and, thus, establish *multi-layer paths* for traffic demands. To speed up the calculation, the heuristic processes traffic demands sequentially.

The consideration of the wavelength assignment problem and non-linear effects which impact the choice of wavelengths for lightpaths is by itself an $\mathcal{NP}$-hard problem [7]. In this paper, we only consider 40 wavelengths per fiber strand which significantly stays behind the possibilities of today's technology (> 160 wavelengths/fiber) to avoid these considerations. Thus, the wavelength assignment problem is assumed to be solved in a separate post-processing step.

B. Related Work

The design of multi-layer networks is a complex issue of high importance for network providers and suppliers. In recent years, network design attracted much attention due to the availability of detailed equipment and CAPEX models [5][8] which base on the IST Nobel project [4].

The authors of [7] use the optical equipment model from [8], introduce an additional "grooming layer", and give ILPs for CAPEX optimization relying on pre-calculated paths to speed up the optimization. For resiliency, 1+1 optical channel protection is considered. In [9], we present ILPs that do not depend on pre-calculated paths and develop heuristics for the considered optimization problem.

The authors of [10] develop ILPs for multi-layer network design based on a detailed theoretical equipment model. They consider 1+1 dedicated protection on IP layer. In [11], they enhance their ILPs by a heuristic branch-and-cut algorithm and focus on two-layer network design. No evaluations are given. In [12], we propose ILPs for CAPEX minimization for transparent, semi-transparent, and opaque optical networks.

The PANEL project [13] considered multi-layer networks with multiple protecting layers. By means of simulations, a quantitative comparison of protection at the highest and lowest layer for ATM and SDH equipment has been performed with a simple CAPEX model. These concepts have been extended in [14] to consider static and dynamic multi-layer recovery strategies by means of simulation studies.

The authors of [15] present estimation formulas for the number of required equipment based on the multi-layer equipment model of [5]. Neither a construction layout nor multi-layer paths for traffic demands are given. In [16], a heuristic for planning GMPLS-based transport networks is presented. The focus is on the optimization of OTN equipment considering wavelengths, wavebands, and fiber.

In [6], we use the CAPEX model [5] and present a CAPEX-aware multi-layer network design algorithm without considering protection. In this paper, we propose a fast deterministic heuristic for CAPEX-aware design of resilient multi-layer networks which provides a construction layout for the multi-layer network and multi-layer paths for the considered traffic demands. The heuristic allows to perform comprehensive parameter studies even on large-scale network instances and traffic matrices in feasible time. We consider protection on the OTN/OCh and IP layer for different topologies.

To the best of our knowledge, a heuristic for CAPEX-aware resilient multi-layer network design for a detailed multi-layer network equipment and CAPEX model providing construction layouts and multi-layer paths has not been presented so far.

IV. HEURISTIC ALGORITHM FOR RESILIENT MULTI-LAYER NETWORK DESIGN

We introduce nomenclature and explain the heuristic for multi-layer network design. Then, we show how the heuristic can be extended for networks with resilience requirements.

A. Nomenclature

The heuristic requires the following input parameters:

- a set $\mathcal{V}$ of *sites / points of presence* (PoP),
- a network equipment and CAPEX model $\mathcal{C}$ which defines the available equipment and a set of layers $\mathcal{L} = \{0, 1, \ldots, 5\}$ where 0 is the physical, 1 the OMS, 2 the OCh, 3 the SDH/SONET, 4 the Ethernet, and 5 the IP/MPLS layer as defined in Section II,
- an unalterable physical topology $\mathcal{G}_0 = (\mathcal{V}, \mathcal{E}_0)$ consisting of a set of directed fiber ducts $\mathcal{E}_0 \subseteq \mathcal{V} \times \mathcal{V}$,
- logical topologies $\mathcal{G}_l = (\mathcal{V}, \mathcal{E}_l)$ with directed edges $\mathcal{E}_l \subseteq \mathcal{V} \times \mathcal{V}$ in layer $l \in \mathcal{L}, l > 0$, and

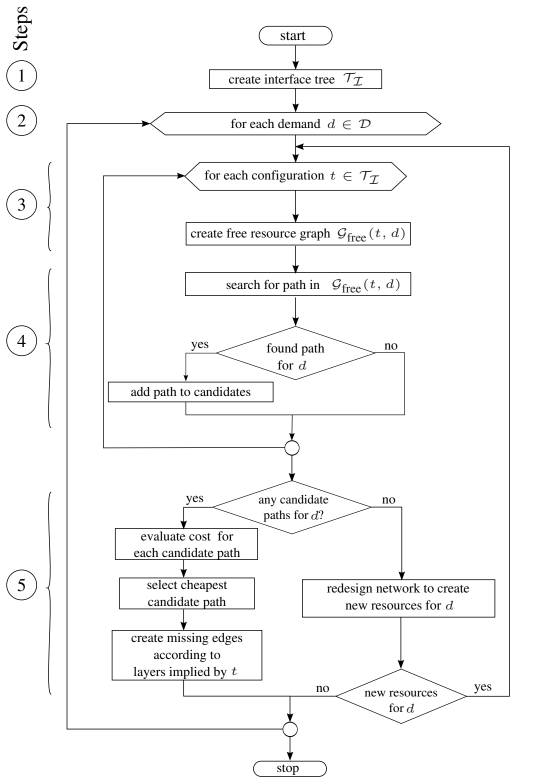


Fig. 5. Workflow of the multi-layer network design heuristic.

- a set $\mathcal{D} \subseteq \mathcal{V} \times \mathcal{V} \times \mathbb{N}$ of directed IP/MPLS traffic demands $d = (s, t, b) \in \mathcal{D}$ with $s, t \in \mathcal{V}$ being d's source and destination site, and $b \in \mathbb{N}$ its bandwidth request in Mbit/s.

From these parameters, we define the following notations:

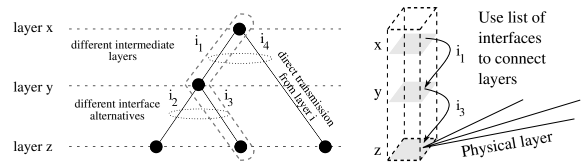
- a set $\mathcal{I} \subset \mathcal{C}$ of tributary and trunk interfaces,
- sets $\mathcal{I}_{xy} \subset \mathcal{I}$ of trunk interfaces connecting layer $x \in \mathcal{L}$ to layer $y \in \mathcal{L}$, $x > y$,
- each edge $e \in \mathcal{E}_l, l \in \mathcal{L}$ is associated with equipment as described in Section II-C and provides a total amount of resources $r_{\text{total}}(e)$ (i.e. capacity or wavelengths) while $r_{\text{free}}(e)$ denotes the amount of its free resources, and
- a function $l : \mathcal{X} \mapsto \mathcal{L}$ indicating the layer of an edge $(\mathcal{X} := \mathcal{E})$ or a component $(\mathcal{X} := \mathcal{C})$, respectively.

In this paper, we consider the layers defined in [4], see Section II-A, and focus on IP/MPLS demands. The heuristic can be easily extended to deal with an arbitrary set of layers $\mathcal{L}$ as well as demands on all available layers.

B. The Algorithm

We give a general overview of the heuristic algorithm for multi-layer network design and following describe the steps of the algorithm in detail.

1) Overview: Fig. 5 illustrates the proposed heuristic algorithm. Initially (step 1), the equipment defined in the CAPEX model $\mathcal{C}$ is analysed to determine the possible interface combinations to connect the IP layer to the physical layer using the interfaces given by $\mathcal{I}$. Following (step 2), the heuristic sequentially processes all demands $d \in \mathcal{D}$ in the given input order and finds or creates multi-layer paths, respectively.



(a) An interface tree listing the possibilities to connect layers x and z via trunk interfaces i_1, i_2, i_3, and i_4

(b) Connecting layers x and z using the highlighted path in the interface tree

Fig. 6. Possibilities to connect multiple layers via interfaces.

For each interface combination that allows to connect the IP to the physical layer (step 3), the heuristic considers the free resources in the traversed layers given by the interfaces. The free resources in an interface combination are used to build a free resource graph consisting of edges with free resources of the traversed layers. A *constrained shortest path first* (CSPF) algorithm is used to search paths which are added to a candidate list for the current demand (step 4).

If there are any candidate paths (step 5), the equipment cost is evaluated to expand the candidate paths of the free resource graph to valid multi-layer paths according to the considered possibility to connect to the physical layer. The cheapest solution is chosen in a greedy fashion and missing equipment is installed. If no path with sufficient free resources exists, the demand will be blocked. To avoid blocking, we consider approaches to alter existing resource allocations such that new resources are created.

These steps are repeated for each demand given by $\mathcal{D}$. Next, the steps of the heuristic are described in detail.

2) Creation of an Interface Tree: The possibilities to connect any layer to any other layer via trunk interfaces can be illustrated by an interface tree. An exemplary interface tree is illustrated in Fig. 6(a) where each edge in the tree is decorated with a trunk interface which connects to a certain lower layer. The tree lists the interfaces $i_1 \in \mathcal{I}_{xy}$, $i_2, i_3 \in \mathcal{I}_{yz}$, $i_4 \in \mathcal{I}_{xz}$. Interfaces i_1 and i_4 differ in the layer they connect to. Interfaces i_2 and i_3 connect to the same layer but might differ, e.g. in the amount of provided capacity. Each path in the tree from the root to a leaf is a possible configuration to connect layers x and z. The path enclosed in dashes is (i_1, i_3).

We define a more general formalism $\mathcal{T}_\mathcal{I}$ which is a set of lists of trunk interfaces that can be used to connect the IP/MPLS layer (5) to the physical layer (0):

$$\begin{aligned} \mathcal{T}_\mathcal{I} := \Big\{ \ (i_1, i_2, \ldots, i_k) \in \mathcal{I}_{x_1 y_1} \times \ldots \times \mathcal{I}_{x_k y_k} : \\ x_1 = 5 \wedge y_1 = x_2 \wedge y_2 = x_3 \wedge \ldots \\ \wedge\, y_{k-1} = x_k, y_k = x_k = 0 \ \Big\} \end{aligned} \tag{2}$$

We map the lists of interfaces given in $\mathcal{T}_\mathcal{I}$ to a set $\mathcal{T}_\mathcal{L}$ of lists of layers that are traversed when a certain interface configuration in $\mathcal{T}_\mathcal{I}$ is used. It is

$$\mathcal{T}_\mathcal{L} := \Big\{ \ \big(l(i_1), \ldots, l(i_k)\big) : (i_1, i_2, \ldots, i_k) \in \mathcal{T}_\mathcal{I} \ \Big\} \,. \tag{3}$$

Thus, each list of interfaces in $\mathcal{T}_\mathcal{I}$ corresponds to a list of layers in $\mathcal{T}_\mathcal{L}$. The combination of $\mathcal{T}_\mathcal{I}$ and $\mathcal{T}_\mathcal{L}$ defines at which layers data is processed and to what extent data is encapsulated and vice versa using interfaces, as depicted in Fig. 6(b). For instance, the path enclosed in dashes in Fig. 6(a) is $(x, y, z) \in \mathcal{T}_\mathcal{L}$. The layer configurations given by $\mathcal{T}_\mathcal{L}$ are used to create free resource graphs, defined next.

3) Selection of Traversed Sites: The selection of the traversed sites for a multi-layer path depends on the existence of paths with sufficient resources.

Therefore, we evaluate all layer configurations given by $\mathcal{T}_\mathcal{L}$ for the processing of a demand $d \in \mathcal{D}$. For each $t = (l_1, \ldots, l_k)$, we define a *free resource graph* $\mathcal{G}_{\text{free}}(t,d) = \big(\mathcal{V}, \mathcal{E}_{\text{free}}(t,d)\big)$ which consists of the edges $\mathcal{E}_{\text{free}}(t,d)$ of all layers listed in t with sufficient free resources for the bandwidth request of d. The construction of $\mathcal{E}_{\text{free}}$ can be formulated as

$$\begin{aligned}\mathcal{E}_{\text{free}}(t,d) := \big\{\ & e \in \mathcal{E}_l : l \in (l_1, \ldots, l_k) = t \in \mathcal{T}_\mathcal{L} \ \wedge \\ & r_{\text{free}}(e) \text{ sufficient for } d\ \big\}.\end{aligned} \tag{4}$$

The layer of the edges implies a rating. For instance, using edges in lower layers is usually cheaper whereas using an edge in the physical layer initially implies high cost and should be avoided, as a rule of thumb. Hence, we apply a weight function $w : \mathcal{E}_{\text{free}} \mapsto \mathbb{R}^+$ on the edges of $\mathcal{G}_{\text{free}}$ which we define as

$$w(e) := \begin{cases} u\big(l(e)\big) \cdot \dfrac{r_{\text{free}}(e)}{r_{\text{total}}(e)} & \text{if } l(e) > 0, \\ 1 & \text{if } l(e) = 0 \wedge e \text{ is unused}, \\ \infty & \text{otherwise}, \end{cases} \tag{5}$$

where the rating function $u : \mathcal{L} \mapsto \mathbb{R}_0^+$ is defined as $u(l) := l/|\mathcal{L}|$ in this paper. Considering the edges' utilization proved a significant enhancement to the AXL algorithm presented in [6]. To find paths in the weighted free resource graph, any k-shortest path algorithm [17] as well as Dijkstra [18] can be used.

4) Decision for a Multi-Layer Path: So far, we have created a weighted free resource graph $\mathcal{G}_{\text{free}}(t,d)$ for each layer configuration $t \in \mathcal{T}_\mathcal{L}$ and used shortest path algorithms to find candidate paths providing sufficient resources for the current demand $d \in \mathcal{D}$.

Since the candidate paths in $\mathcal{G}_{\text{free}}(t,d)$ contain edges from any layer given by the layer configuration $t \in \mathcal{T}_\mathcal{L}$, the candidate paths have to be extended with further equipment to be suited for the traffic demand's requirements. In particular, we have to ensure connectivity of d's source and destination on the IP layer as we consider IP/MPLS demands in this paper.

We calculate the equipment that is needed to complete the candidate paths according to the corresponding interface configuration $\mathcal{T}_\mathcal{I}$ and determine the additional cost using a CAPEX function $c : \mathcal{C} \mapsto \mathbb{R}_0^+$ given by $\mathcal{C}$ that associates cost to the network equipment.

Finally, we select the cheapest of the candidates in a *greedy* fashion and install the required equipment extensions.

5) Network Resource Redesign: If no path with sufficient free resources can be found, a demand would be blocked. To avoid blocking, we consider approaches to alter existing resource allocations such that new resources are established to carry further demands. We denote these approaches by *network resource redesign*. An optimal network-wide redesign is a complex task as it requires that all established traffic demands are considered for de- and reallocation of resources. Hence, our heuristic focuses on the redesign of single edges in logical layers and applies the following two mechanisms where appropriate.

TABLE I
DETAILS ON THE CONSIDERED NETWORK TOPOLOGIES [19].

Topology	$\lvert\mathcal{V}\rvert$	$\lvert\mathcal{E}_0\rvert$	$\lvert\mathcal{D}\rvert$	Total IP/MPLS traffic (Gbit/s)	Average bandwidth / demand (Gbit/s)
Nobel Germany (G17)	17	52	121	665.6	5.9
Germany50 (G50)	50	176	662	2344.9	3.5
Nobel Europe (EU)	28	82	378	1943.6	5.1
Nobel U.S. (US)	14	42	91	5550.1	61.0

a) Interface upgrades: The number of resources provided by an interface increases with its price. Initially, cheap interfaces providing few resources are installed which can be replaced by more expensive and more powerful interfaces if more resources are needed.

b) Aggregation at intermediate layers: An additional transport layer provides additional resources but also causes additional cost. Therefore, its usage should be avoided until additional resources are required. For instance, there is no need to use the *optical multiplex section* (OMS) layer as long as only few demands are routed in a network. When the number of demands increases, more resources are needed which can be realized by using the *wavelength division multiplex* (WDM) capabilities of the OMS layer. Existing logical edges are rerouted to make use of the intermediate transport layer. The equipment associated with these edges is changed accordingly.

Both approaches create new free resources which are available in the free resource graph and can be used to aggregate further edges to mitigate resource blocking and increase resource utilization.

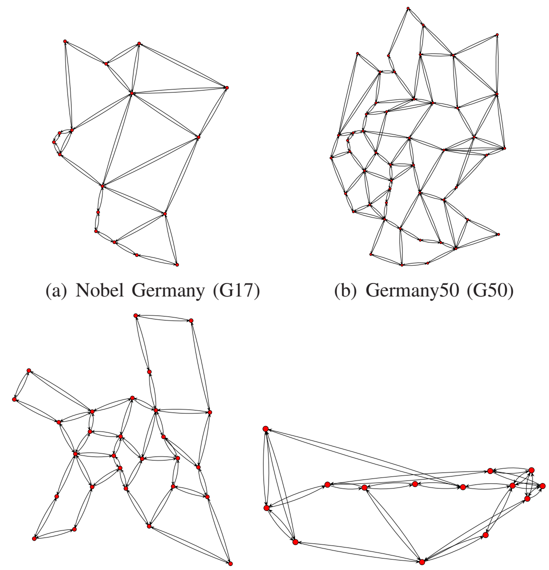
C. Protection

Compared to [6], we also extended the heuristic to realize dedicated protection for each demand on a given layer with different resilience requirements. Primary and backup paths for a traffic demand are searched alternatingly, i.e. the backup path is searched after the primary path. During the search of the primary and backup paths as well as network redesign operations, the disjointness of primary and backup according to the considered resilience mechanisms must be preserved.

The heuristic sequentially processes demands in the input order. Due to this demand-wise process, it is possible to apply a different resilience mechanism for each demand in general.

V. EVALUATION

In this section, we perform evaluations for four different resilience requirements on four physical network topologies. All evaluations have been run on a Linux-based PC with a 2.4 GHz processor using a Java application [20] developed by the authors which implements the proposed heuristic.



(a) Nobel Germany (G17) (b) Germany50 (G50)

(c) Nobel Europe (EU) (d) Nobel U.S. (US)

Fig. 7. THE CONSIDERED TOPOLOGIES [19] WITH DIRECTED EDGES.

The considered resilience requirements are the unprotected case which serves as a reference, *link protection* (LP) on OCh layer for fast local recovery, and *path protection* (PP) on IP layer either with *link disjointness* (LD) or *node disjointness* (ND) for disaster recovery. In this work, we focus on the protection on a single layer and the same resilience mechanism is used for all demands. The considered physical network topologies are taken from [19], include an IP traffic matrix $\mathcal{D}$, and have directed links. *Nobel Germany* (G17) is an exemplary German backbone topology with 17 sites which are a subset of the sites used in *Germany50* (G50) which is a finer grained German backbone network with 50 sites. *Nobel Europe* (EU) is a European reference network that was also used in *COST266*. *Nobel U.S.* (US) is a reference long-range transport network whose topology is also known as *NSFNET*. Numerical details for all four topologies are listed in Table I. Illustrations of the four physical topologies are shown in Fig. 7.

A. CAPEX and Resource Utilization

We performed evaluations to assess the performance of the heuristic for all combinations of the considered network topologies and resilience requirements. The numerical results of these evaluations are listed in Table II.

In the second column of Table II, the number of routed demands is listed. The number of total demands $|\mathcal{D}|$ for a network scenario is given in Table I. The IST Nobel equipment model [4] defines interfaces up to 40 Gbit/s. Without modifications to the traffic matrices, all demands requesting more than 40 Gbit/s would be blocked. Therefore, we split all such demands into as many 40 Gbit/s demands as possible and fill the rest up with demands of 10 Gbit/s. The EU topology contains one such demand from Glasgow to London at 54 Gbit/s which increases the number of demands from 378 to 380. As a pure long-range transport network, the US topology is intended to carry demands with very high requests for bandwidth. The highest request in this data set is around 330 Gbit/s. After the split-up of these demands, there are 225 demands in total and the average capacity per demand decreases to 24.7 Gbit/s. There are no demands above 40 Gbit/s in G17 and G50. Blocking of demands only occurs in the US scenario with LP on OCh layer which contains rather few links and is not suited for LP.

TABLE II
EVALUATIONS OF TOPOLOGIES AND RESILIENCE REQUIREMENTS.

Topology & resilience	Routed	Protected	CAPEX	IP util.	Comp. time (s)
G17 unprot.	121	0	2510.54	0.72	3
G17 OCh LP	121	119	3800.77	0.66	4
G17 IP PP LD	121	112	3644.95	0.55	4
G17 IP PP ND	121	78	3427.91	0.51	4
G50 unprot.	662	0	9648.79	0.80	28
G50 OCh LP	662	657	16075.15	0.80	98
G50 IP PP LD	662	654	14399.41	0.66	81
G50 IP PP ND	662	602	14188.76	0.68	116
EU unprot.	380	0	6948.28	0.80	8
EU OCh LP	380	369	12779.20	0.79	27
EU IP PP LD	380	367	9727.90	0.67	21
EU IP PP ND	380	286	9629.66	0.65	27
US unprot.	225	0	10783.22	0.81	1
US OCh LP	211	171	15288.55	0.80	5
US IP PP LD	225	223	15427.38	0.59	3
US IP PP ND	225	222	15711.53	0.56	5

The third column of Table II shows the number of demands that were not only routed but also protected according to the considered resilience requirement. If no backup path can be found for a demand, only the primary path is set up, i.e. the backup path is blocked. We can see that the number of routed and protected demands behaves similarly for all four topologies with each resilience requirement, except for US with LP on OCh layer. Protection on OCh layer with LP is able to protect the most demands – except for US – as the setup of lightpaths in the OCh layer can make direct use of the underlying OMS layer which is able to provide many resources and degrees of freedom. Protection on the IP layer is not able to find as many backup paths as OCh protection. In case of IP with PP and ND, less demands can be protected compared to mere LD as possibilities to find a backup path which shares no node with the primary path are drastically reduced. The decrease depends on the overall degrees of freedom provided by the underlying physical topology.

The fourth column of Table II shows the overall CAPEX for the multi-layer network according to [4]. The resulting CAPEX depends on the number of routable and protectable demands. As with the resilience mechanism, a similar behavior can be seen for all considered network topologies regarding CAPEX. Exemplarily, the results for the G50 topology are illustrated in Fig. 8 and Fig. 9 and are explained in the following. Fig. 8 shows that the overall CAPEX for a network with dedicated protection is not twice the cost for the unprotected network although at least twice the resources

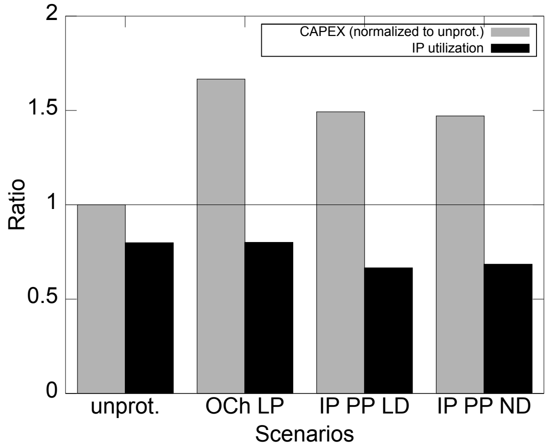


Fig. 8. CAPEX AND IP UTILIZATION IN THE G50 SCENARIOS.

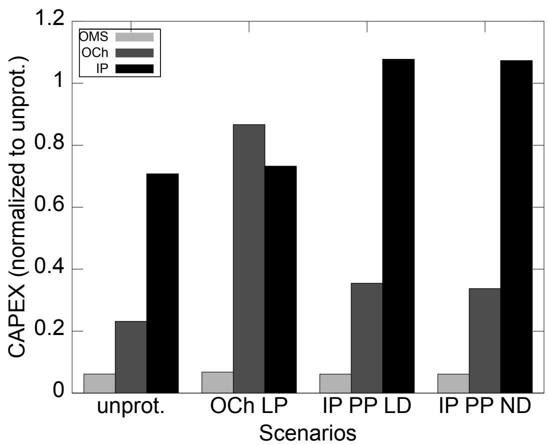


Fig. 9. CAPEX SPLIT UP BY LAYERS IN THE G50 SCENARIOS.

are required to route the primary and backup demands. To compare the impact of the different resilience requirements on the CAPEX value, we not only look at the overall CAPEX, but also at the individual CAPEX of the used layers. This split-up into OMS, OCh, and IP layer CAPEX is illustrated in Fig. 9 for the G50 scenarios. The physical layer is not listed here as we only consider zero-cost components with it. As shown in previous work [6], IP equipment is expensive. IP is the main cost factor for all resilience scenarios except LP on OCh layer. In the latter case, the IP topology is nearly identical to the unprotected IP topology which results in almost identical IP equipment cost. However, the equipment that has to be installed in the OCh layer to realize link protection is drastically increased. This can be also seen in the physical path lengths for these scenarios depicted in Fig. 10 which shows the minimum, average, and maximum lengths of primary and backup paths. In case of LP on OCh, the path lengths are almost three times higher on average than with the other resilience mechanisms. Therefore, the overall CAPEX for the LP on OCh scenarios is always the highest, except for the US topology which provides relatively few links at all.

The fifth column of Table II lists the capacity utilization on the IP layer, i.e. the ratio of used to installed capacity in

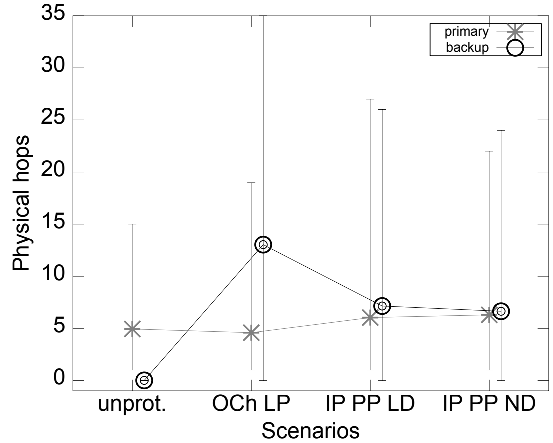


Fig. 10. PATH LENGTHS IN THE PHYSICAL LAYER OF THE G50 SCENARIOS.

the IP layer. As the heuristic intends to increase the resource utilization such that the amount of installed equipment as well as CAPEX is reduced, the IP utilization can be seen as an indicator for the goodness of the proposed heuristic's results. The IP utilization is lower with protection as additional disjointness constraints must met. Especially, the utilization for protection on IP layer is lower as disjoint routings must be found in the IP layer itself, i.e. additional IP equipment has to be installed. However, the IP utilization also slightly decreases for protection on the OCh layer. This might not seem to be intuitive as no additional IP equipment has to be installed. The decrease results from the sequential processing of the demands which can lengthen the primary path of a demand which is also shown in Fig. 10.

The last column of Table II shows the computation time of the evaluations in seconds. As the heuristic processes demands sequentially, the computation time of the heuristic is significantly improved compared to the complexity estimation shown in Equation (1). Each evaluation was completed within less than two minutes. This outlines the feasibility of the proposed heuristic for comprehensive parameter studies.

B. Failure and SRG Analysis

Due to the recursive realization of logical edges in lower layers, multi-layer networks imply *shared risk groups* (SRG), e.g. when a fiber link fails, all logical links using this fiber fail subsequently. Therefore, we analyze the impact of certain failures regarding the size of the SRGs and their impact regarding affected demands. We consider all *single-link* (SL) failures and all *single-node* (SN) failures. Multi-homing to save demands in case their source or destination site fails was not applied in this paper. The results for the G50 scenarios are illustrated in Fig. 11 for both failure types.

In Fig. 11, the average number of failed links is illustrated. For SL failures in the physical layer, this number equals to one. For SN failures, it is the sum of the average incident node degree on all layers of the failing site. Fig. 11 shows that LP causes a higher node degree and is not suited for node failures.

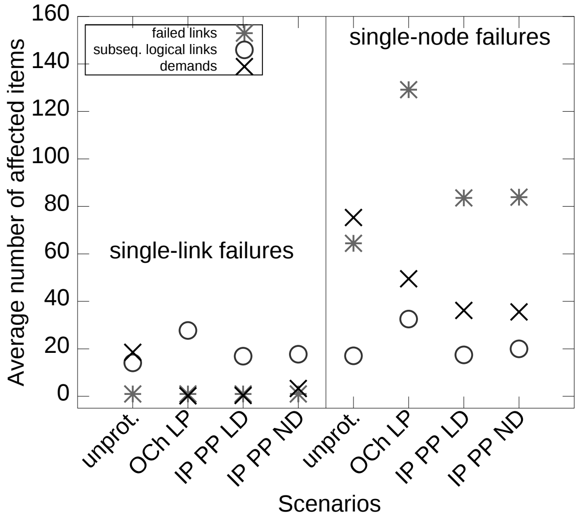


Fig. 11. FAILURE ANALYSIS IN THE G50 SCENARIOS.

We also investigate the average number of subsequently failed logical links, i.e. the size of SRGs. With LP on the OCh layer the number of subsequently failed links is higher, both for SL and SN failures as much more links in OCh layer exist than in the PP scenarios. For PP with ND, the average SRG size is slightly higher than for PP with LD as more links are set up for PP with ND to bypass sites in the backup path that are used in the primary path.

In the unprotected case, the average number of affected demands is highest as no backup paths are set up at all. With SL failures, no demand fails with any of the considered resilience strategies. With SN failures, the PP on IP layer results in less affected demands than the LP in OCh layer as LP is not suited for node failures at all.

VI. SUMMARY

In this paper, we presented a heuristic for cost-aware resilient multi-layer network design using a detailed network equipment model to provide a construction layout for a multi-layer network. We performed evaluations for four reference network scenarios with up to 50 sites. The heuristic proved to yield results with a high resource utilization. No evaluation took more than two minutes which outlines the feasibility of the proposed heuristic for comprehensive parameter studies.

We discussed the impact of different resilience requirements on the network's *capital expenditure* (CAPEX), blocking probabilities, path lengths in the physical layer for primary and backup paths. Furthermore, we analyzed the amount of affected traffic with the different resilience and network scenarios as well as the size of *shared risk groups* (SRG) in case of failures.

In future work, we will compare the results of our heuristic to mathematical optimization approaches using *integer linear programming* (ILP), study multi-layer resilience, and consider shared protection to further reduce cost.

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] "ITU-T Recommendation G.709 (2009), Interfaces for the Optical Transport Network (OTN)," ITU, Dec. 2009.

[3] M. Pióro and D. Medhi, *Routing, Flow, and Capacity Design in Communication and Computer Networks*. Morgan Kaufman, 2004.

[4] IST NOBEL Working Group, "NOBEL and NOBEL-2 Project Website," Apr. 2010, http://www.ist-nobel.org.

[5] 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.

[6] 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.

[7] M. Scheffel, R. G. Prinz, C. G. Gruber, A. Autenrieth, and D. A. Schupke, "Optimal Routing and Grooming for Multilayer Networks with Transponders and Muxponders," in *IEEE Globecom* , San Francisco, CA, USA, Nov. 2006.

[8] M. Gunkel, R. Leppla, M. Wade, A. Lord, D. A. Schupke, G. Lehmann, C. Fürst, S. Bodamer, B. Bollenz, H. Haunstein, H. Nakajima, and J. Martensson, "A Cost Model for the WDM Layer," in *Photonics in Switching*, Creete, Greece, Oct. 2006.

[9] M. Duelli, C. Pluntke, and M. Menth, "Minimizing Installation Costs of Survivable DWDM-Mesh Networks: A Heuristic Approach," in *4th Euro-NGI Conference on Next Generation Internet Networks (NGI 2008)*, Krakow, Poland, Apr. 2008, pp. 15–22.

[10] S. Orlowski and R. Wessäly, "An Integer Programming Model for Multi-Layer Network Design," Konrad-Zuse-Zentrum für Informationstechnik Berlin, ZIB Preprint ZR-04-49, Dec. 2004. [Online]. Available: http://www.zib.de/PaperWeb/abstracts/ZR-04-49

[11] S. Orlowski, A. M. Koster, C. Raack, and R. Wessäly, "Two-layer Network Design by Branch-and-Cut featuring MIP-based Heuristics," Konrad-Zuse-Zentrum für Informationstechnik Berlin, Tech. Rep., 2006.

[12] C. Pluntke, M. Menth, and M. Duelli, "CAPEX-Aware Design of Survivable DWDM Mesh Networks," in *IEEE International Conference on Communications (ICC)*, Dresden, Germany, Jun. 2009.

[13] P. Demeester, M. Gryseels, A. Autenrieth, C. Brianza, L. Castagna, G. Signorelli, R. Clemente, M. Ravera, A. Jajszczyk, D. Janukowicz, K. V. Doorselaere, and Y. Harada, "Resilience in Multilayer Networks," *IEEE Communications Magazine*, vol. 37, no. 8, pp. 70–76, Aug. 1999.

[14] M. Pickavet, P. Demeester, D. Colle, D. Staessens, B. Puype, L. Depré, and I. Lievens, "Recovery in Multilayer Optical Networks," *Journal of Lightwave Technology*, vol. 24, no. 1, p. 122, 2006.

[15] C. Pavan, R. M. Morais, and A. N. Pinto, "Estimating CapEx in Optical Multilayer Networks," in *Conference on Telecommunications (ConfTel)*, May 2009, pp. 335–338.

[16] N. Naas and H. T. Mouftah, "Cost-effective heuristics for planning GMPLS-based transport networks," *Optical Switching and Networking*, vol. 7, no. 1, pp. 12–27, 2010.

[17] A. W. Brander and M. C. Sinclair, "A comparative study of k-shortest path algorithms," in *Proc. 11th UK Performance Engineering Worksh. for Computer and Telecommunications Systems*, Sep. 1995.

[18] E. W. Dijkstra, "A Note on Two Problems in Connexion with Graphs," *Numerische Mathematik*, vol. 1, pp. 269–271, 1959.

[19] 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.

[20] *100GET-E3-R3G-2L: A Tool for Visualizing and Evaluating Multi-Layer Networks*, Apr. 2010, http://www3.informatik.uni-wuerzburg.de/100GET.