

# Network Resilience through Multi-Topology Routing

Michael Menth, Ruediger Martin  
Department of Distributed Systems,  
Institute of Computer Science,  
University of Würzburg,  
Am Hubland, 97074 Würzburg, Germany  
E-mail: {menth, martin}@informatik.uni-wuerzburg.de

**Abstract**—In this paper we propose the use of multi-topology (MT) routing for network resilience against link and node failures. We describe the multi-topologies by an  $n$ -dimensional vector of different link costs for all links in the network. It is the base for the calculation of  $n$  shortest path trees from any node to all other destinations, i.e. for  $n$  virtual routing topologies. We define the link costs in such a way that the routing topologies complement each other in the sense that at least one valid route remains in a single link or node failure scenario for each pair of nodes in at least one routing topology. In such a failure case, packets are rather deviated over the intact routing topology than discarded. The recovery speed of the presented mechanism is very fast and can be compared to fast rerouting mechanisms in MPLS, which reduce packet drops to a minimum. In contrast to MPLS, MT routing is still a pure IP-based solution that retains the scalability and the robustness of IP routing.

## I. INTRODUCTION

Critical operations such as tele-monitoring and control, or telesurgery depend on very reliable communication regarding packet loss, delay, or connectivity. Thus, carrier grade networks must offer a high availability of 99.999%, which is also called the “five nines” property. Service interruptions should be avoided and their duration should be kept to a minimum if they occur. Hence, fault tolerance and resilience of a network becomes a QoS feature because outages can significantly disturb the user experienced network QoS.

Regarding network failures IP networks are self-healing, i.e., connectivity is restored automatically after link state information about a failure has been propagated through the network. This is a very robust mechanism but it requires a substantial amount of time that cannot be afforded by demanding applications. Therefore, fast reroute mechanisms, e.g. based on MPLS, are defined that try to minimize the recovery time in such a case. An IP solution of this problem is the subject of this paper.

Multi-topology (MT) routing provides several different IP routing topologies within one network. It is an optional mechanism within IS-IS ([1], C.2) used today by many Internet

service providers (ISPs) for Interior Gateway Protocol (IGP) routing within their clouds [2]. MT routing can be used for a variety of purposes such as an in-band management network “on top” of the original IGP topology, to maintain separate IGP routing domains for isolated multicast or IPv6 islands within the backbone, or to force a subset of an address space to follow a different topology. Recently, MT routing extensions for OSPF were proposed that offer the maintenance of different link costs for each virtual multi-topology [3], [4]. We take advantage of this new mechanism to provide a new robust mechanism for resilient routing.

We keep the mechanism simple. One MT routing scheme is used under normal networking conditions. If a node detects the outage of one of its adjacent links or neighbor nodes, it deviates all traffic that has to be sent according to the routing table over this failed element to another interface, over an alternative route that is provided by a another MT routing scheme.

The paper is organized as follows. In Section 2 we describe existing resilience mechanisms and MT routing as proposed by the IETF. In Section 3 we propose the use of MT routing for fast rerouting for locally recognized failures. We illustrate its operation, clarify logical requirements, and compare it with existing resilience mechanisms. Section 4 discusses open issues and Section 5 summarizes this work.

## II. RELATED WORK

In this section we give a short overview on related resilience mechanisms for which MT routing could be a reasonable alternative and contrast them to other resilience mechanisms with a different focus. Then, we explain MT routing as defined by the IETF.

### A. Resilience Mechanisms

IP networks have the self-healing property, i.e., their routing re-converges after a network failure by exchanging link state vectors such that all but the failed nodes can be reached after a while – if a working path still exists. MPLS technology provides resilience mechanisms that set up backup tunnels in advance such that the traffic may be simply switched to the backup path if the primary path fails. This is called protection switching as opposed to path restoration in IP [5].

This work was funded by the Bundesministerium für Bildung und Forschung of the Federal Republic of Germany (Förderkenzeichen 01AK045) and Siemens AG, Munich. The authors alone are responsible for the content of the paper.

1) *Economic Rerouting*: The backup paths require additional resources. The 1+1 protection approach sends the traffic simultaneously over the primary and the backup path. The 1:1 protection approach uses the backup path only if the primary path fails such that backup path resources may be shared by different backup paths that are activated in different failure scenarios. Resource sharing is the key for economic resilience. It can be exploited more systematically by MPLS than by pure IP solutions due to explicit routing of label switched paths (LSPs) as opposed to shortest path routing. However, attempts are already made to reduce the required backup resources for IP networks [6]. Multi-path routing together with load balancing can further reduce the amount of required backup capacity such that 17% additional capacity suffices to protect all single link failures in the COST239 network (cf. Figure 2(a)) [7].

2) *Fast Rerouting*: IS-IS and OSPF are the most widely used interior gateway routing protocols (IGPs). They are link state protocols, i.e., the routers exchange link state advertisements (LSAs) in regular intervals and calculate the shortest path trees based on this structure. The interval length is in the order of 10 seconds and cannot be reduced to arbitrarily small values [8]. In addition, the computation of the shortest paths that are needed to construct the routing tables is based on the new LSAs and requires a substantial amount of time. This time overhead is tolerable for elastic traffic but not for real-time traffic or even high-precision telematic or tele-surgery applications.

Protection switching is faster as the backup paths already exist. However, if a primary path fails, the outage may be somewhere along the path. Therefore, the outages have to be signalled to the ingress router where the backup paths starts. This procedure takes some time. Thus, fast rerouting requires that a backup path is available at the location where the outage occurs [9], i.e., a primary path needs a backup path starting at each of its nodes. These so-called one-to-one backups may reconnect to the primary paths further downstream or they may directly lead to the destination. The one-to-one backup option entails obviously tremendously many paths and states inside the network. This is reduced by the so-called facility backup option where each link or node failure is protected by separate backup paths. Thus, a facility backup deviates many primary paths at once around the failure locations such that they can reconnect their primary paths again. However, this still induces a substantial amount of backup paths and the resulting number of states puts a great burden on the involved routers. Fast rerouting with 1:1 protection requires substantially more capacity than economic rerouting but in contrast to 1+1 protection, the additional capacity may be used to carry low priority traffic in failure-free scenarios.

## B. Multi-Topology Routing

We first explain multi-topology (MT) routing according to IETF and explain then our use of that concept.

1) *MT Routing for IS-IS*: The authors of [2] proposed MT routing extensions for IS-IS.

We denote a network as a directed graph  $\mathcal{G} = (\mathcal{V}, \mathcal{E}, k(\cdot))$ . The set of vertices  $\mathcal{V}$  represents the routers and the set of directed edges  $\mathcal{E}$  represents the links in the network. Traditional routing protocols require a cost function  $k(l)$  for the links  $l \in \mathcal{E}$ . In case of link state routing protocols, each node broadcasts the link costs in regular intervals or in case of topology changes such that any node in the network has a complete map about active links and their costs. Based on this information, each node computes a shortest paths tree [10] and determines the next hop for each destination within the network which is recorded - possibly in a compacted form - in the routing table. In the following, we describe the multi-topology (MT) routing approach currently described by the IETF and our proposal.

MT routing provides  $n$  different routing schemes  $R_i$  in a network that are characterized by their unique MT ID # $i$  with  $0 \leq i < n$ . The current Internet draft proposes to create different  $R_i$  by including links to the MT or by excluding them. Hence, a new virtual network topology  $\mathcal{G}_i = (\mathcal{V}_i, \mathcal{E}_i, k(\cdot))$  is created that differs in  $\mathcal{E}$ . For backward compatibility reasons,  $\mathcal{E}_0 = \mathcal{E}$  contains all links in the network. By omitting some links in the topology, the broadcast of link state packages (LSPs) is limited to the nodes within the same MT and the shortest path algorithm is performed for each  $\mathcal{G}_i$  to calculate a separate routing table  $T_i$  for each  $R_i$ . Normal data packets are marked with one MT ID. If a packet is received by a forwarding process, the MT ID # $i$  of the packet is evaluated first and then the next hop towards the destination of the packet is derived as usual from routing table  $T_i$ . This way, a subset of the traffic in the network can be forced to use only the link subset  $\mathcal{E}_i \subseteq \mathcal{E}$ .

2) *MT Routing for OSPF*: In [3], [4] similar extensions are proposed for OSPF. It enhances the scalar link costs to  $n$ -dimensional vectors  $\mathbf{k}(l) = \begin{pmatrix} k(l)_0 \\ \vdots \\ k(l)_{n-1} \end{pmatrix}$  where  $k(l)_i$  corresponds to routing scheme  $R_i$ . Hence, the routing protocol now exchanges MT-specific link costs. In contrast to the above scheme, all nodes and links may participate in all routing topologies  $\mathcal{G}_i$  that differ only in their cost function  $k(l)_i$ . The shortest path computation is executed for each topology  $\mathcal{G}_i$  and produces a separate routing table  $T_i$  for each routing scheme  $R_i$  due to the different link costs. Like above, packets are marked with the number of their MT routing and forwarded according to the respective routing tables.

Setting the MT specific link costs to a very high value has the same effect as excluding the link from the network topology. We say that those links are not contained in the routing topology although they are present in  $\mathcal{E}_i$  but they are usually not used for packet forwarding within this topology. Hence, this MT routing scheme is at least as powerful as the one above. The advantage of the new concept for OSPF is that every routing topology retains the full self-healing potential of IP routing since all physical links  $\mathcal{E}$  are qualified to repair the connectivity of a routing scheme  $R_i$ . Hence, after some time,  $R_i$  is up again and may continue to forward packets. In

particular, connectivity can be restored even in large outage scenarios as long as working paths exist within the topology. This is a significant advantage over MPLS solutions. Note that the original scheme for IS-IS is limited to the subset of links  $\mathcal{E}_i \subseteq \mathcal{E}$  which reduces the self-healing potential.

### III. NETWORK RESILIENCE THROUGH MT ROUTING

In this section, we explain how MT routing for OSPF can enhance the resilience of IP networks. We discuss the requirement regarding the virtual routing topologies and compare it with existing rerouting mechanisms.

#### A. Illustration of MT Routing Resilience

The idea is the following. Packets are forwarded according to a routing scheme  $R_i$  and if a link or router failure occurs in  $R_i$ , the MT ID of the affected traffic is changed locally by the router that detects the failure and has problems to forward the packet. For that purpose, the routing table  $T_i$  is enhanced by a backup routing scheme such that the new MT ID can be looked up and inserted into the packet header. The packet is then forwarded according to the routing table of the new MT ID. If the new MT is also broken somewhere in the network, this can possibly create loops. Therefore, an additional time-to-live (MT TTL) is required which is initially set to the maximum MT ID and decreased whenever a MT change occurs. If the MT TTL is zero, the packet is discarded. This method works well both with our new MT routing concept and the original proposal by the IETF.

We illustrate this concept based on the artificial example network in Figure 1(a). We define the MT routing schemes  $R_i$  in such a way that packets are routed on the spanning trees depicted in Figures 1(b)–1(d). Table I shows the routing tables  $T_0^A$ ,  $T_1^A$ , and  $T_2^A$  of router A for the respective MT routing schemes. For each destination router, it provides the corresponding output interface  $IF_i^A$ . In addition, the backup routing topology is indicated. For example, a packet at router A with MT ID #0 and destined for router B is forwarded over  $IF_0^A$ , i.e., it is sent to link A–B. If A–B is down, its MT ID is changed to #1 and the MT TTL is decremented. Then, the corresponding output link is A–C. If this link is also down, the MT ID is changed either to #0 or to #2 depending on the result of a hash algorithm that is applied to the packet header. Hence, load balancing can be applied, but it should be done on the flow level [11], [12]. The MT TTL is again decremented. Assuming that both link A–B and link A–C fail, the MT ID of the packet is changed in a circle until MT TTL is zero and the packet is eventually discarded.

The new concept for the implementation of MT routing (OSPF) allows that a MT routing scheme may be self-healing like conventional IP routing. It takes a while until the new network topology information has been exchanged after a failure and until the routing tables  $T_i$  of the affected routing schemes  $R_i$  have been set up again. If it is necessary to maintain the connectivity within a single routing topology, links with large costs are integrated into the virtual routing topology, thus, increasing the respective set of used edges  $\mathcal{U}$ .

#### B. Requirements for Resilient MT Routing

The resilience mechanism above requires at least two different topologies such that any network node can reach any other destination by a working routing topology if an adjacent link or node fails. Thus, in this section we explain the requirements for the virtual routing topologies such that the different routing schemes protect each other and we give examples that show the existence of virtual topologies in real networks which satisfy these conditions.

1) *Virtual Topology Requirements*: The set of used edges  $\mathcal{U}_i$  of a routing topology  $R_i$  comprises those links that contribute to a shortest path according to the link costs in a failure-free networking scenario. If a link  $l$  fails, any routing topology containing  $l$  is possibly corrupted. Hence, MT routing can be resilient to the failure of a link  $l$  if at least one routing topology exists that does not contain  $l$ . If  $v$  is an interior node in the graph of used links of a routing topology, it may serve as transit router and its failure could corrupt the routing scheme. Hence, MT routing can be resilient to the failure of a node  $v$  if at least one routing topology exists where  $v$  is a leaf node in its graph of used links.

2) *Examples*: Figures 2(a) and 2(b) show the physical topologies of the core of the COST-239 testbed [13] and the testbed of the KING project (Key Components for the Internet of the Next Generation, [14]). In both networks, dual routing schemes  $R_0$  and  $R_1$  with the properties discussed above are easy to find.

Figures 3(a) and 3(b) illustrate two dual virtual routing topologies for each of the two networks that protect each other against single link or node failures. According to the requirements, any node is leaf node in at least one routing topology and any link is not contained in at least one routing topology.

Note that the routing topologies need not be spanning trees. Adding the dashed links destroys the spanning tree structure of the virtual topologies but does not violate the resilience constraints. Apart from that, not all physical links can be included into the routing topologies as they would violate the resilience constraints concerning node failures for at least one of the both routing topologies.

The concept of “resilient routing layers” (RRL) has been presented in [15], [16]. The layers correspond to our virtual routing topologies but they are not based on MTs. Hence, a single layer is not self-healing as it can not be extended by additional links if network elements fail. This, however, is necessary to maintain connectivity in unplanned failure scenarios. The authors of [15], [16] have also proposed algorithms to construct suitable RRLs. An adaptation of these methods can possibly help to generate also virtual routing topologies automatically.

#### C. Comparison with Existing Rerouting Mechanisms

MT routing is clearly a fast rerouting mechanism. In contrast to MPLS fast rerouting it reveals the following advantages. Two virtual topologies are enough to guarantee end-to-end reachability with resilience whereas MPLS solutions require

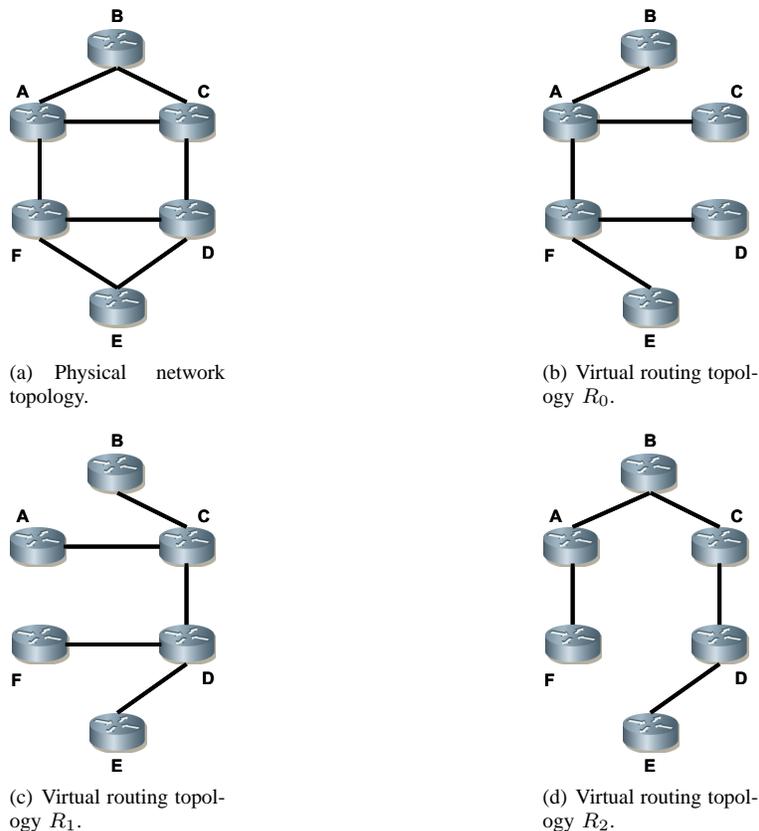


Fig. 1. The physical network topology and multiple virtual routing topologies for a small example network.

TABLE I  
ROUTING TABLES  $T_0^A$ ,  $T_1^A$ , AND  $T_2^A$  OF ROUTER A.

destination	$IF_0^A$	backup for $R_0$	$IF_1^A$	backup for $R_1$	$IF_2^A$	backup for $R_2$
B	A-B	$R_1$ (100%)	A-C	$R_0$ (50%), $R_2$ (50%)	A-B	$R_1$ (100%)
C	A-C	$R_2$ (100%)	A-C	$R_2$ (100%)	A-B	$R_1$ (100%)
D	A-F	$R_1$ (50%), $R_2$ (50%)	A-C	$R_0$ (50%), $R_2$ (50%)	A-B	$R_1$ (100%)
E	A-F	$R_1$ (50%), $R_2$ (50%)	A-C	$R_0$ (50%), $R_2$ (50%)	A-B	$R_1$ (100%)
F	A-F	$R_1$ (100%)	A-C	$R_0$ (50%), $R_2$ (50%)	A-F	$R_1$ (100%)

$|\mathcal{V}| \cdot (|\mathcal{V}| - 1)$  primary paths for the mere end-to-end reachability and additional backup paths for all links in the primary paths (one-to-one backup). This is a severe problem for label switching routers whose maximum number of simultaneous label switched paths (LSPs) is limited. MPLS fast rerouting does not protect against double failures because connectivity is lost if both primary and backup path break. In such a case, MT routing maintains at least connectivity because simple IP rerouting for each MT topology is performed which also considers links with expensive costs which are outside the virtual routing topology. Hence, MT routing is as robust as IP routing and more resilient than MPLS fast rerouting in case of unplanned multiple failures.

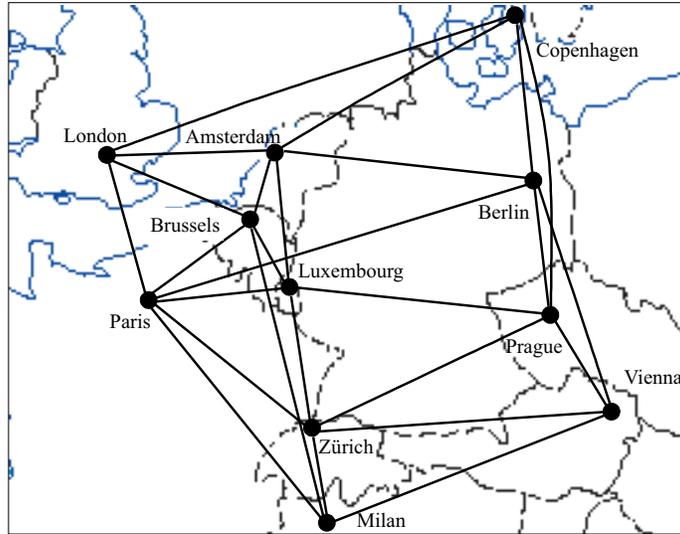
Now, we consider the resource requirements for resilient MT routing. We may use two different virtual routing topologies, one for primary transport and one for backup transport. This corresponds to maintaining two separate networks from

a resource point of view. Therefore, the double amount of capacity is required than for a network without resilience features. In addition, topologies with fewer links lead to longer paths, therefore, more than 100% backup capacity is needed, which is in the same order of magnitude like the requirements for MPLS fast rerouting. Hence, MT rerouting is not economic.

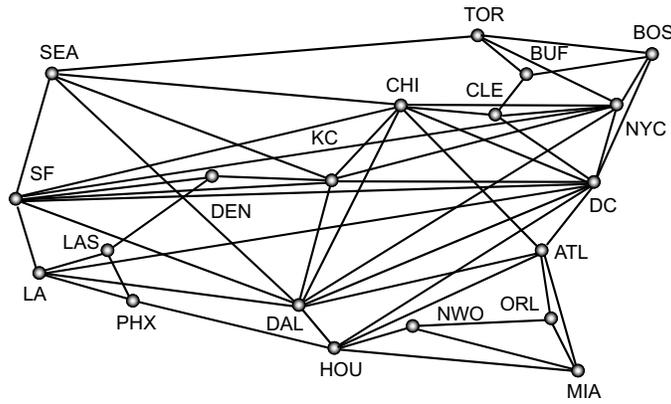
#### IV. OUTLOOK ON FURTHER RESEARCH

So far, we have made plausible that MT routing is an attractive means to implement fault tolerance in IP networks. We identify the following items as open research issues.

- *Resilient Network Dimensioning*. In case of network outage, QoS in terms of loss and delay can only be guaranteed if the bandwidth is sufficient for the backup traffic. Therefore, network failures and corresponding backup capacities must be taken into account in the network dimensioning process.

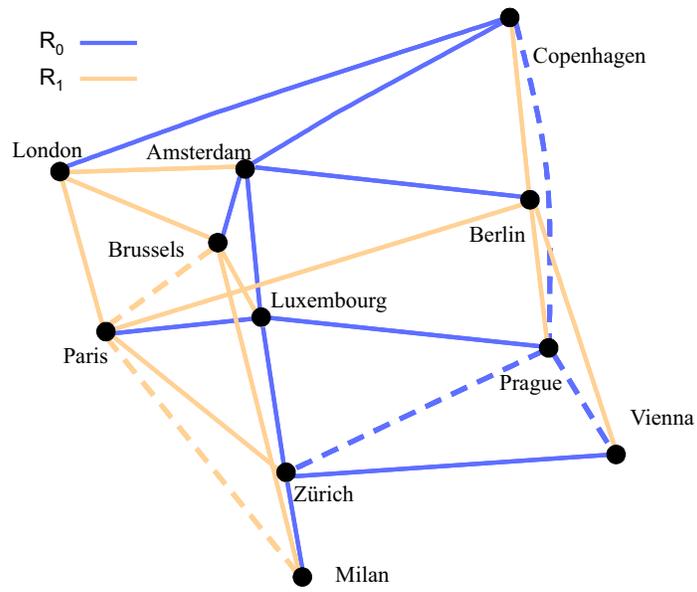


(a) COST-239network.

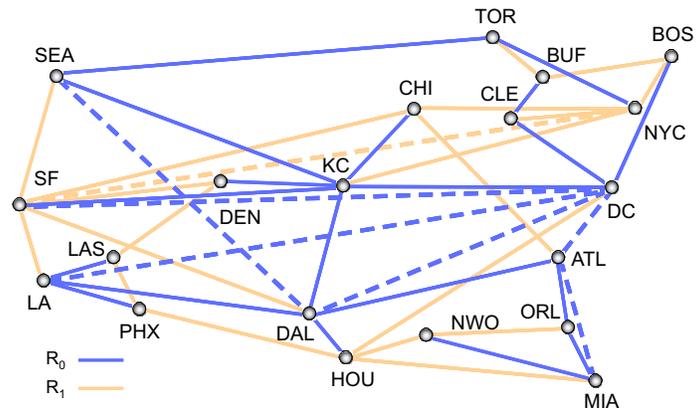


(b) Labnet03 network.

Fig. 2. Physical network topologies.



(a) COST-239 network.



(b) Labnet03 network.

Fig. 3. Dual routing topologies protect against all single link and node failures.

- *Consideration of Shared Risk Link Groups (SRLGs).* In this paper we did not consider links with shared risks, i.e., links that fail simultaneously because they depend on the same fiber or duct. We are aware of that problem which has to be respected for the dimensioning of networks (regarding suitable locations for backup capacity), for the assignment of link costs, or for the primary and backup path definition, respectively. SRLGs introduce a high computational complexity [17] such that these problems are often solved by heuristic algorithms [18], [19].
- *Minimization of Propagation Delay.* The virtual routing topologies should be designed in a way that the mere propagation delay is not excessive due to an excessive length of the shortest paths offered by the default routing topology. The use of several routing topologies helps if the topologies with the shortest paths for a specific aggregate are used for primary and backup routing. Note that the logical requirements in Section III-B are intentionally formulated in such a way that they are applicable to more than two complementing virtual routing topologies.
- *Minimization of Capacity Requirements.* The needed network capacity for the backup routes can be reduced by sharing backup capacities intentionally, i.e. by 1:1 protection. That means, several routing topologies are required with overlapping backup routes such that backup capacity can be saved on common links. In addition, multi-path routing may be applied to further reduce the capacity requirements, similarly to the concept of Self-Protecting Multi-Path [7]. Well designed systems can achieve considerable bandwidth savings without losing network resilience.
- *Automation of Link Cost Assignment.* The link cost assignment for MT routing should be automated by offline tools or by distributed algorithms to construct resilient MT topologies. Optimal solutions and fast heuristics are welcome.
- *Determination of the Complexity of the above Optimization Problems.* The assignment of virtual routing topologies with the above mentioned constraints and optimization goals poses interesting computational problems whose complexity is not investigated yet.

We briefly assess the importance of these issues. The SRLG issue must be solved to make resilient MT routing applicable, which can be done by heuristics. The network dimensioning is straightforward if the routing scheme and the set of failures to protect is given. The minimization of the propagation delay and the capacity requirements is optional and need not be solved to use resilient MT routing. The automation of the link cost assignment is desirable for ease of network management. The determination of the problem complexity leads to interesting insights regarding the adaptability of resilient MT routing and justifies the use of suboptimal heuristic algorithms.

## V. SUMMARY

In this paper we reviewed existing rerouting strategies. They are either economic or fast. We explained multi-topology

(MT) routing and enhanced this mechanism for fast rerouting. We illustrated the operation of the new concept and derived the logical requirements of the complementing virtual topologies. We showed its viability in two example topologies and compared resilient MT routing with existing resilience mechanisms. Finally, we discussed open issues that provide interesting aspects for further research.

## ACKNOWLEDGEMENTS

The authors would like to thank Prof. Tran-Gia for the stimulating environment at the Department of Distributed Systems and Dr. Charzinski for the fruitful discussions.

## REFERENCES

- [1] Callon, R.: RFC1195: Use of OSI IS-IS for Routing in TCP/IP and Dual Environments. <http://www.ietf.org/rfc/rfc1195.txt> (1990)
- [2] Przygienda, T., Shen, N., Sheth, N.: M-ISIS: Multi Topology (MT) Routing in IS-IS. <http://www.netzmafia.de/rfc/internet-drafts/draft-ietf-isis-wg-multi-topology-07.txt> (2004)
- [3] Psenak, P., Mirtorabi, S., Roy, A., Nguyen, L., Pillay-Esnault, P.: MT-OSPF: Multi Topology (MT) Routing in OSPF. <http://ietfreport.isoc.org/ids/draft-ietf-ospf-mt-04.txt> (2005)
- [4] Mirtorabi, S., Roy, A.: Multi-Topology Routing in OSPFv3 (MT-OSPFv3). <http://ietfreport.isoc.org/ids/draft-mirtorabi-mt-ospfv3-02.txt> (2005)
- [5] Autenrieth, A., Kirstädter, A.: Engineering End-to-End IP Resilience Using Resilience-Differentiated QoS. *IEEE Communications Magazine* **40** (2002) 50–57
- [6] Nucci, A., Schroeder, B., Bhattacharyya, S., Taft, N., Diot, C.: IGP Link Weight Assignment for Transient Link Failures. In: 18<sup>th</sup> International Teletraffic Congress (ITC), Berlin (2003)
- [7] Menth, M.: Efficient Admission Control and Routing in Resilient Communication Networks. PhD thesis, University of Würzburg, Faculty of Computer Science, Am Hubland (2004)
- [8] Iannaccone, G., Chuah, C.N., Bhattacharyya, S., Diot, C.: Feasibility of IP Restoration in a Tier-1 Backbone. *IEEE Network Magazine (Special Issue on Protection, Restoration and Disaster Recovery)* (2004)
- [9] Pan, P., Swallow, G., Atlas, A.: Fast Reroute Extensions to RSVP-TE for LSP Tunnels. <http://www.ietf.org/internet-drafts/draft-ietf-mpls-rsvp-lsp-fastreroute-07.txt> (2004)
- [10] Dijkstra, E.W.: A Note on Two Problems in Connexion with Graphs. *Numerische Mathematik* **1** (1959) 269 – 271
- [11] Cao, Z., Wang, Z., Zegura, E.: Performance of Hashing-Based Schemes for Internet Load Balancing. In: *IEEE Infocom, Tel Aviv, Israel* (2000)
- [12] Dittmann, G., Herkersdorf, A.: Network Processor Load Balancing for High-Speed Links. In: *International Symposium on Performance Evaluation of Computer and Telecommunication Systems (SPECTS)*, San Diego, CA (2002) 727–735
- [13] Batchelor et al., P.: Ultra High Capacity Optical Transmission Networks. *Final report of Action COST 239* (1999)
- [14] Hoogendoorn, C., Schrodi, K., Huber, M., Winkler, C., Charzinski, J.: Towards Carrier-Grade Next Generation Networks. In: *International Conference on Communication Technology (ICCT)*, Beijing, China (2003)
- [15] Cicic, T., Hansen, A.F., Gjessing, S., Lysne, O.: Applicability of Resilient Routing Layers for k-Fault Network Recovery. In: *International Conference on Networking (ICN)*, Reunion Island (2005)
- [16] Kvalbein, A., Hansen, A.F., Cicic, T., Gjessing, S., Lysne, O.: Fast Recovery from Link Failures Using Resilient Routing Layers. In: *IEEE Symposium on Computers and Communications (ISCC)*, Cartagena, Spain (2005)
- [17] Hu, J.Q.: Diverse Routing in Optical Mesh Networks. *IEEE/ACM Transactions on Networking* **51** (2003) 489 – 494
- [18] Zhan, H., Ou, C.S., Mukherjee, B.: Path-Protection Routing and Wavelength Assignment (RWA) in WDM Mesh Networks under Duct-Layer Constraints. *IEEE/ACM Transactions on Networking* **11** (2003) 248–258
- [19] Todimala, A., Ramamurthy, B.: IMSH: An Iterative Heuristic for SRLG-Diverse Routing in WDM Mesh Networks. In: *IEEE International Conference on Computer Communications and Networks (ICCCN)*, Chicago, IL (2004)