University of Würzburg
Institute of Computer Science
Research Report Series

Network Resilience through Multi-Topology Routing

Michael Menth, Ruediger Martin

Report No. 335 May 2004

University of Wuerzburg, Germany
Institute of Computer Science
Department of Distributed Systems
Am Hubland, D-97074 Würzburg, Germany
phone: (+49) 931-8886644, fax: (+49) 931-8886632
{menth|martin}@informatik.uni-wuerzburg.de

Network Resilience through Multi-Topology Routing

Michael Menth, Ruediger Martin
University of Wuerzburg, Germany
Institute of Computer Science
Department of Distributed Systems
Am Hubland, D-97074 Würzburg, Germany
phone: (+49) 931-8886644, fax: (+49) 931-8886632
{menth|martin}@informatik.uni-wuerzburg.de

Abstract

In this paper we propose the use of multi-topology (MT) routing to achieve fault tolerance against failures of network elements which is also called network resilience. The shortest path routing trees seen from any node is calculated based on the link costs. To provide different routing topologies, an n-dimensional vector of different costs is provided for each link in the network and allows for the calculation of n different routing topologies. We suggest to set these vectors in such a way that any node has a valid path to any other destination in case of a link or router failure. If a router recognizes the outage of an interface, it can switch the traffic from a broken routing topology to a valid routing topology which leads to a very fast reaction time to failures. The failover time can be compared with MPLS solutions that are based on explicit routing. However, we change the implementation of MT routing also in such a way that the new concept is still based on the shortest path convergence mechanism. This offers the potential to maintain the robustness of the Internet routing in case of multiple failures.

Keywords: Resilience Methods, IP Rerouting

1 Introduction

In this paper we do not consider traditional traffic engineering approaches to achieve high bandwidth utilization or to avoid quality of service (QoS) violations in terms of packet loss an delay. Instead, we consider resilience against network failure as performance measure because carrier grade networks are required to offer a high availability and reliability of 99.999%, which is also called the "five nines" property. Hence, service interruptions should be avoided and their duration should be kept to a minimum if they occur.

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

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

providers (ISPs) for Interior Gateway Protocol (IGP) routing within their clouds [2]. MT-routing can be used for variety of purposes such as an in-band management network "on top" of the original IGP topology, maintain separate IGP routing domains for isolated multicast or IPv6 islands within the backbone, or force a subset of an address space to follow a different topology.

In this paper we enhance MT routing to provide fault tolerance against network failures. The idea is simple. One distinct MT routing scheme is used under normal normal networking conditions. If a node detects the outage of one of its adjacent links or neighbor node, it deviates all traffic, that has to be sent according to the routing table over this failed element, to another interface over an alternative routed provided by a another MT routing scheme.

If a failure happens in the physical network topology, the link or node vanishes also in the virtual routing topology for the respective MT routing. Each of these topologies retains the self-healing property of IP routing although this may take some time until the routing tables have converged. We change the implementation of MT routing proposed in [2] in such a way that each routing scheme is based on the full physical network topology in order to take any link for restoration purposes. This increases the self-healing property of a single MT routing scheme.

The paper is organized as follows. In Section 2 we describe MT routing in detail and propose our implementation which has an increased potential for self-healing. Section 3 enhances MT routing conceptually to allow for fast rerouting in case locally recognized failures. In Section 4 we derive virtual topology requirements for resilient routing and illustrate their feasibility in sample networks. Section 5 summarizes this work and gives some outlook on further research.

2 Multi-Topology Routing

We denote a network as a directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, k(\dot{)})$. 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 [3] 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 create different R_i by including links to the MT or to exclude them. Hence, a new virtual network topology $\mathcal{G}_i = (\mathcal{V}_i, \mathcal{E}_i, k(\dot{)})$ 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 calculated 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 $\mathcal{E}_i \subseteq \mathcal{E}$.

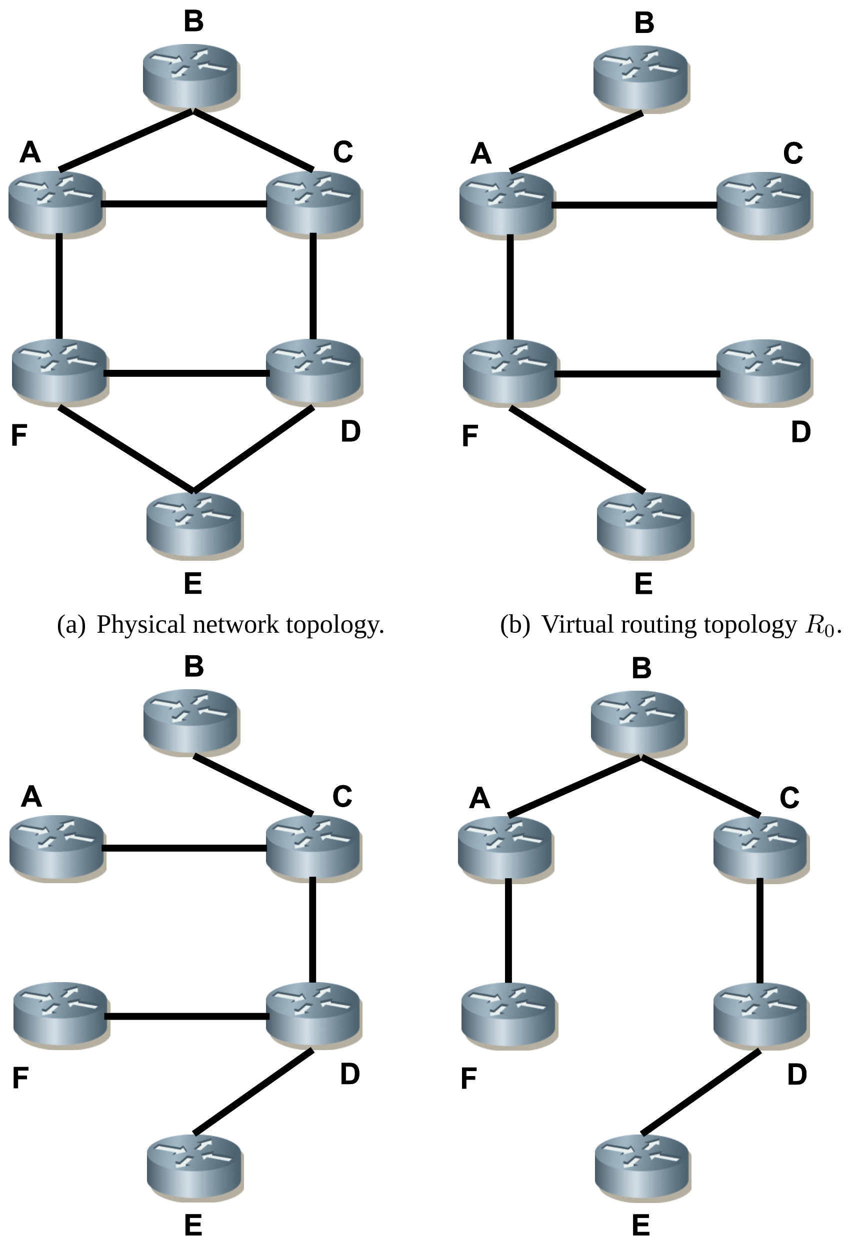
For our purposes we use another approach to provide different routing schemes R_i, which is not yet being standardized by the IETF. The scalar link costs are enhanced to n-dimensional vectors $\mathbf{k}(\mathbf{l}) = \begin{pmatrix} k(l)_0 \\ \cdot \\ k(l)_{n-1} \end{pmatrix}$ and $k(l)_i$ corresponds to routing scheme R_i. The routing protocol requires now the broadcast of link state vectors. In contrast to the above scheme, all nodes an links participate in all network 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 implementation is that every routing topology can basically use all physical links $\mathcal{E}$ to repair the connectivity of a routing scheme some time after a failure. In contrast, the original scheme is limited to the subset $\mathcal{E}_i \subseteq \mathcal{E}$ for that purpose.

3 Network Resilience through MT Routing

In this section, we explain how MT routing can enhance the resilience of IP networks. 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. To that aim, 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 1 shows the routing tables T_0^A, T_1^A, and T_2^A of router A for the respective MT routing schemes. The next hop for routing R_i is determined by the interface IF_i^A. If this interface fails, the MT-ID of the packets is changed to a backup MT-ID. At this occasion, load balancing can be applied, however, it should be done on the flow level [4, 5]. Assuming that both link A-B and A-C fails, the MT-ID of the packet is changed in a circle until MT-TTL is zero and packet is eventually discarded.



(a) Physical network topology.

(b) Virtual routing topology R_0.

(c) Virtual routing topology R_1.

(d) Virtual routing topology R_2.

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

Table 1: 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-F	R_1 (100%)
F	A-F	R_1 (100%)	A-C	R_0 (50%), R_2 (50%)	A-F	R_1 (100%)

The new concept for the implementation of MT routing allows that a MT routing scheme may be self-healing like conventional IP routing. It takes a while until the network topology 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.

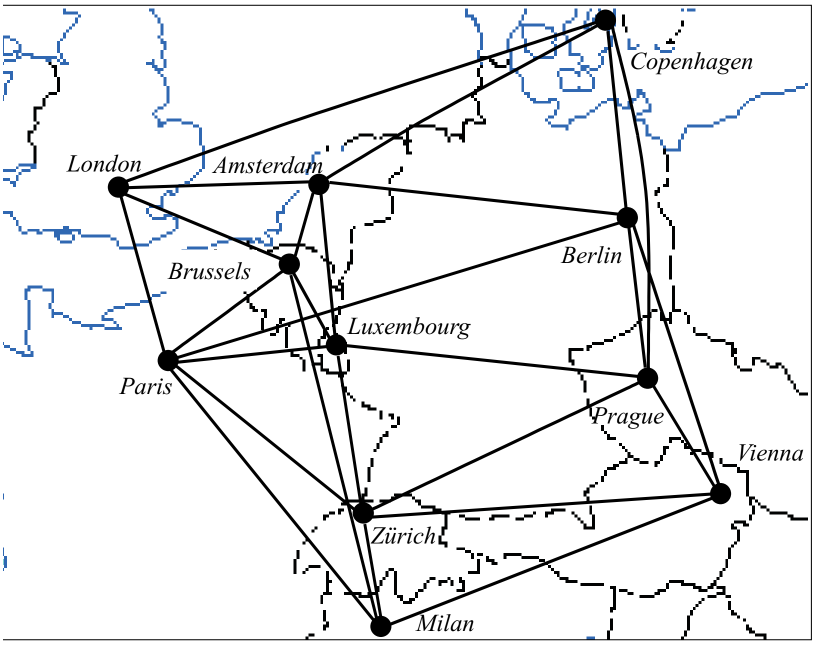
Our proposal increases the reaction speed upon a network failure while the enhancement of the MT routing offers some potential to make this concept also robust against multiple network failures.

4 Configuration of MT Routing

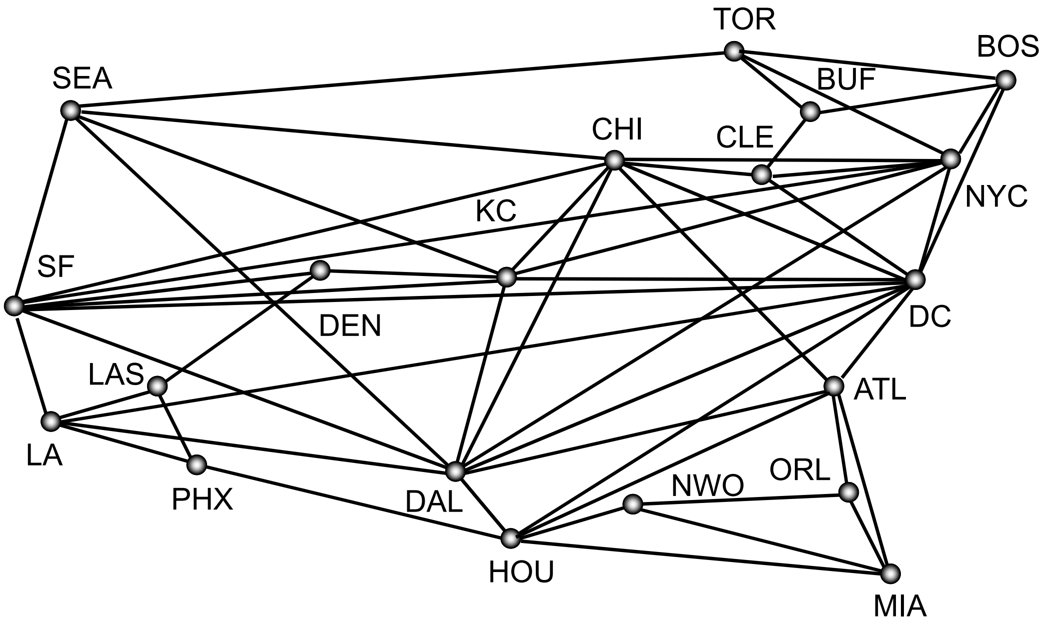
The mechanism in the previous section postulates that any node in the network is able to find a valid routing topology if a neighboring link or node fails. This in this section we explain the requirements for the virtual routing topologies such that the different routing schemes protect each other. Finally, we show that it is possible to fulfill these requirements in real networks.

If a link l fails, any routing topology containing l in its routing topology is possibly corrupted. Hence, MT routing can be resilient to the failure of a link l if there is at least one routing topology that does not contain l.

If a node v is a leaf node within a routing topology, it can fail without hampering the remaining routing scheme since only the traffic destined for v itself can not be routed anymore to its destination. Conversely, if v is an interior node within a routing topology, it may serve as transit router and its failure would corrupt the routing scheme. Hence, MT routing can be resilient to the failure of a node v if there is at least one routing topology where v is a leaf node.

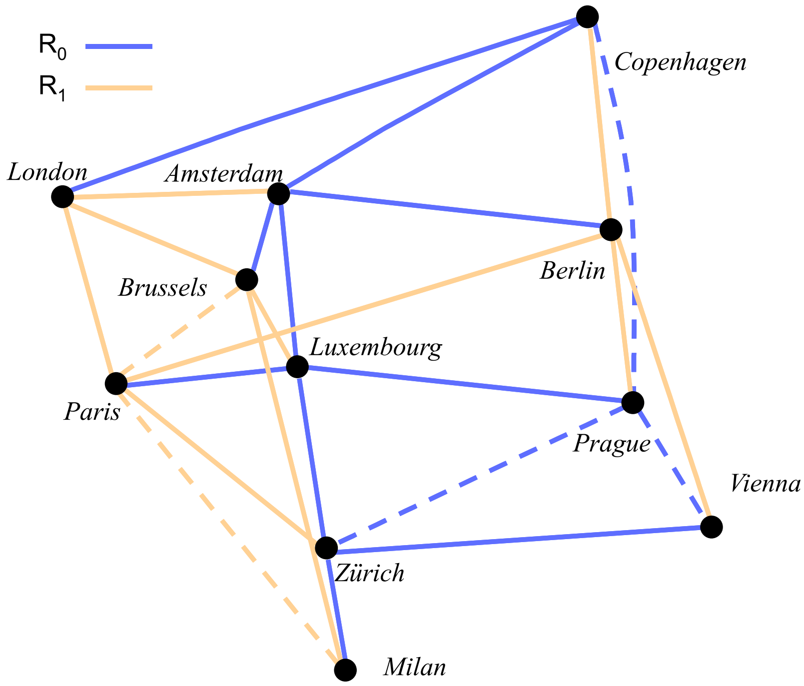


(a) COST-239network.

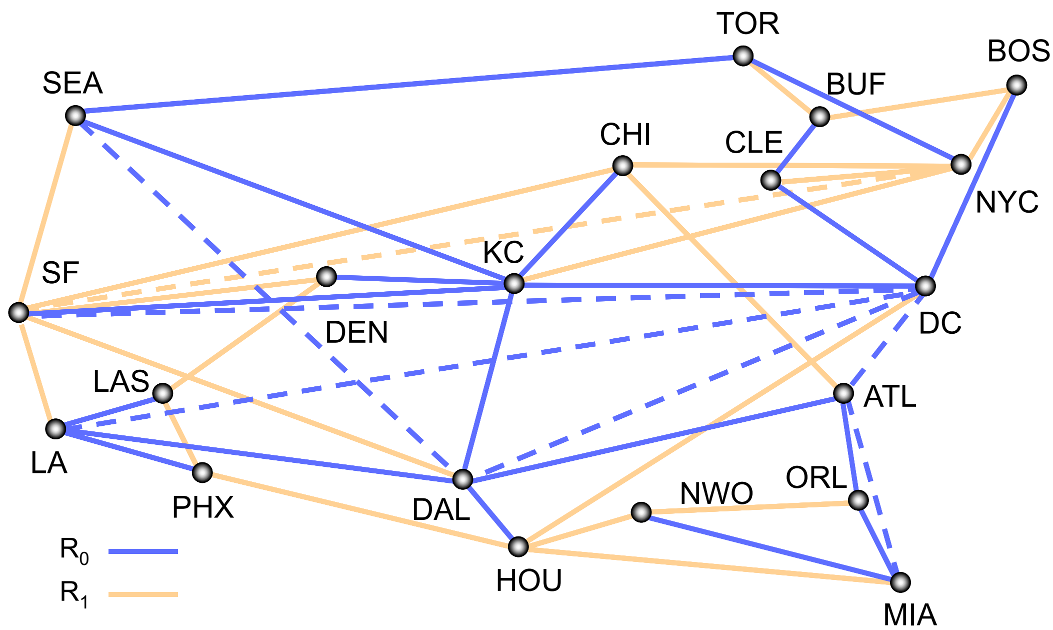


(b) Labnet03 network.

Figure 2: Physical network topologies.



(a) COST-239 network.



(b) Labnet03 network.

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

Figure 2(a) shows the physical topology of the core network of the COST-239 testbed [6] and Figure 2(b) shows the physical topology of the Labnet03 which is the testbed of the KING network [7]. In both networks it is easy to find dual routing schemes R_0 and R_1 such that they protect the network against all link and node failures, i.e., any node is leaf node in at least one routing topology and any link is not contained in at least one routing topology. Figures 3(a) and 3(b) show the corresponding dual routing topologies. Note that the routing topologies do not need to be spanning trees. The dashed lines can be added without violating the protection criterion but they destroy the spanning tree structure. Apart from that, not all physical links can be included into the routing topologies as they would violate the criterion for resilience against all node failures for at least one of the both routing topologies.

5 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 point.

- The cost assignment of the links should be automated by offline tools and by distributed algorithms to construct resilient MT topologies.

- In case of network outage, QoS in terms of loss and delay can only be optimized if the bandwidth is sufficient for the deviated traffic. This poses two different optimization problems. First, the routing and the link capacities should be designed such that the required backup capacity is minimized to support a given traffic matrix and a set of protected failure scenarios $\mathcal{S}$ [8]. This is an approach for network dimensioning. Second, for an existing network with a given traffic matrix the routing should be designed that the maximum link utilization is as small as possible in any protected failure scenario [9]. Load balancing deviated traffic onto different backup topologies can improve the results but a side condition for both challenges is keeping the number of different MT routing schemes small. In addition, load balancing must be performed on the flow level [4, 5].

- The transient behavior of MT routing must be studied if the configuration of the different topologies is changed.

- The transient behavior of MT routing must be studied and enhancements have to be made if more than a single network failure occurs.

- In general, MT routing does not always use shortest paths and increases thereby the transmission delay, in particular, if a failure occurs. The routing topologies might be optimized regarding this aspect.

- The application of MT routing is probably most interesting in Ad-Hoc networks to improve the routing stability. We intend to investigate this issue and to adapt typical Ad-Hoc routing algorithms to MT routing.

6 Summary

In this paper we have described an alternative implementation for multi-topology (MT) routing and a new concept to achieve fast rerouting through MT routing. We have illustrated this resilience mechanism by a small example. We have derived conditions such that resilience against all link and node failures can be provided by using several MT routing topologies. We have applied this concept in two example networks an showed that this mechanism is well applicable and that two different routing topologies suffice to achieve full resilience against all single link and node failures. Finally, we have pointed out various research aspects regarding network resilience through MT routing.

Acknowledgements

The authors would like to thank Prof. Tran-Gia for the stimulating environment at the Department of Distributed Systems, Dr. Charzinski for the fruitful discussions, and Dr. Vu Phan-Gia for the preparation of the routing schemes in the example networks.

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://vesuvio.ipv6.cselt.it/internet-drafts/draft-ietf-isis-wg-multi-topology-06.txt (2003)

[3] Disjkstra, E.W.: A Note on Two Problems in Connexion with Graphs. Numerische Mathematik **1** (1959) 269 – 271

[4] Cao, Z., Wang, Z., Zegura, E.: Performance of Hashing-Based Schemes for Internet Load Balancing. In: IEEE INFOCOM'00, Tel Aviv, Israel (2000)

[5] Dittmann, G., Herkersdorf, A.: Network Processor Load Balancing for High–Speed Links. In: SPECTS 2002, San Diego, CA (2002) 727–735

[6] Batchelor et al., P.: Ultra High Capacity Optical Transmission Networks. Final report of Action COST 239. http://barolo.ita.hsr.ch/cost239/network/ (1999)

[7] Hoogendoorn, C., Schrodi, K., Huber, M., Winkler, C., Charzinski, J.: Towards Carrier-Grade Next Generation Networks. In: ICCT 2003, Beijing, China (2003)

[8] Menth, M., Reifert, A., Milbrandt, J.: Self-Protecting Multipaths - A Simple and Resource-Efficient Protection Switching Mechanism for MPLS Networks. In: Networking 2004, Athens, Greece (2004)

[9] Nucci, A., Schroeder, B., Bhattacharyya, S., Taft, N., Diot, C.: IGP Link Weight Assignment for Transient Link Failures. In: 18th International Teletraffic Congress (ITC18), Berlin (2003)