University of Würzburg
Institute of Computer Science
Research Report Series

The Effect of Combining Loop-Free Alternates and Not-Via Addresses in IP Fast Reroute

Rüdiger Martin[1], Michael Menth[1], Matthias Hartmann[1],
Tarik Čičić[2], Amund Kvalbein[2]

Report No. 432 September 2007

[1] University of Würzburg
Institute of Computer Science
Chair of Distributed Systems
Würzburg, Germany
Email: {martin|menth|hartmann}@informatik.uni-wuerzburg.de

[2] Simula Research Laboratory
Oslo, Norway
Email: {tarikc|amundk}@simula.no

The Effect of Combining Loop-Free Alternates and Not-Via Addresses in IP Fast Reroute

Rüdiger Martin, Michael Menth, Matthias Hartmann
University of Würzburg, Institute of Computer Science
Am Hubland, D-97074 Würzburg, Germany
Email: {martin|menth|hartmann}@informatik.uni-wuerzburg.de

Tarik Čičić, Amund Kvalbein
Simula Research Laboratory
Oslo, Norway
Email: {tarikc|amundk}@simula.no

***Abstract*—The IETF currently discusses fast reroute mechanisms for IP networks (IP FRR) to accelerate the recovery in case of network element failures and to avoid microloops during network-wide routing re-convergence. Several mechanisms are proposed. Loop-free alternates (LFAs) are simple, but they cannot cover all single link and node failures. Not-via addresses are more complex and cover all single failures, but they potentially lead to longer backup paths and require tunnelling which may reduce the forwarding speed of the routers. In addition, they increase the size of the forwarding tables.**

This work studies the combination of those simple and complex mechanisms to achieve full single failure coverage with least overhead. First, we establish a taxonomy for LFAs according to their ability and propose combination options with not-vias for different resilience requirements. Then, we quantify the effect of combining both mechanisms regarding their applicability for the resilience requirements, routing table size, link utilization, backup path length, and amount of traffic requiring decapsulation per router. The results show that there are no strong advantages of the combined application of both mechanisms over not-via as the only IP FRR concept if 100% coverage for single link and node failures is required.

I. Introduction

Given the growing size and complexity of modern communication networks, network element failures are a fact of their daily operation [1] and require special precautions. To that end, resilience mechanisms maintain connectivity in case of outages where possible.

Resilience mechanisms can be divided into restoration and protection schemes. Restoration sets up a new path after a failure while protection switching pre-establishes backup paths in advance. IP rerouting implements restoration. It is robust [2], [3], but slow: although careful tuning of timeout parameters reduces the recovery time to values in the order of one second [4], [5], this time cannot be reduced arbitrarily without jeopardizing the network stability [5]. In contrast, multiprotocol label switching (MPLS) technology has the ability to implement protection switching by pre-establishing explicitly routed backup paths in advance. The primary and backup path concept requires the activation of the backup path by some form of failure notification, but fast reaction times in the order of 100 ms can be achieved.

New emerging services such as Voice over IP, virtual private networks for finance, and other real-time business applications require stringent service availability and reliability. Their demand for a very fast reaction to failures lead to the development of fast reroute (FRR) techniques. That means, backup paths are not only available at the source of a primary paths but at each intermediate node of a path for immediate local reaction. For MPLS, two different FRR approaches have already been standardized [6]. However, pure IP networks also need fast resilience. Therefore, current IETF drafts and other publications propose various methods for IP FRR [7]–[11].

IP FRR is also designed to prevent packet loss caused by micro-loops during the routing re-convergence of IP networks. Local failure recovery suppresses network-wide failure notification and thereby global re-convergence. This avoids microloops for short-lived failures which is a big advantage since 50% of all failures last less than a minute [1], [12]. In case of long-lived failures, IP FRR is useful to gain time for ordered loop-free convergence as suggested in [13].

Besides, the mechanisms should be simple and deployable in the current routing architecture. They should cover most failures, e.g., all single link or node failures, and they should not create problems, e.g. unpredictable severe routing loops, in case of unanticipated multiple failures.

In this context we focus on the IP FRR mechanisms that are currently discussed in the IETF. In case of failures, loop-free alternates (LFA) redirect traffic to neighboring nodes having a shortest path towards the destination avoiding the failed element [8]. Not-via addresses provide local IP-in-IP tunnels to the next-next-hop (NNHOP) to bypass the failed element [9]. LFAs are simple as they avoid tunnels and they potentially lead to shorter detours, but they cannot protect all single failures. Some LFAs are able to protect only link failures, others protect also router failures. Some lead to routing loops in case of multiple failures, others are safe. Not-via addresses are more complex as new prefixes need to be distributed via routing protocols. They require tunnelling which is undesirable as decapsulation potentially reduces the forwarding speed of the routers and might lead to packet fragmentation due to MTU limitations. However, not-via addresses offer 100% failure coverage.

The contribution of this paper is twofold. Firstly, we provide a classification of different LFAs with respect to their ability and establish a new taxonomy. Secondly, we study the effect of combining appropriate LFAs and not-via addresses to achieve 100% coverage. We discuss the pros and cons of both mecha-

nisms and analyze their applicability for different resilience requirements. We also discuss the backup path length, the impact on routing table size, the resource utilization in terms of link load, and the amount of tunnelled traffic.

The paper is structured as follows. Section II introduces a new taxonomy of LFAs according to their ability. Section III explains the concept of not-via addresses. In Section IV we discuss pros and cons of both mechanisms and propose several combination options thereof for different resilience requirements. Section V presents and interprets the results of our experimental analysis. After a short discussion of related work in Section VI, we conclude this work in Section VII.

II. Classification of Loop-Free Alternates

In this section, we review the definition of LFAs, we classify them according to their ability, and establish a new taxonomy.

A. Definition of LFAs

A loop free alternate (LFA) is a local alternative path from a source node S towards a destination D in the event of a failure [8]. If S cannot reach anymore its primary next hop P towards D, it simply sends the traffic to another neighbor N that still can forward the traffic to D avoiding both the failed element and S and thus does not create routing loops. LFAs are pre-computed and installed in the forwarding information base of a router for each destination. The Internet draft [8] specifies three criteria for LFAs to guarantee different levels of protection quality and loop avoidance. We illustrate these conditions and provide a taxonomy to classify neighbor nodes with respect to their ability to be used as LFAs.

B. Loop-Free Condition (LFC)

We consider source S and destination D in Fig. (1). The numbers associated with the links are the link metrics taken into account for shortest path routing. When the link $S \rightarrow P$ fails, packets can only be rerouted over neighbor N. However, this creates a forwarding loop because the shortest path of N to D leads over S. Therefore, N cannot be used as LFA by S to protect the failure of link $S \rightarrow P$. To avoid loops, the following loop-free condition (LFC) must be met:

$$\text{dist}(N, D) < \text{dist}(N, S) + \text{dist}(S, D). \quad (1)$$

In Fig. (2) both neighbors N_1 and N_2 of source S fulfill this condition with regard to destination D.

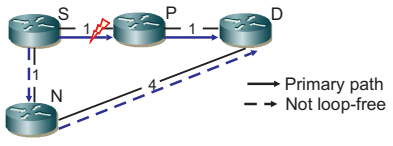


Fig. 1: The neighbor N cannot be used as LFA because it does not meet the loop-free condition (LFC).

C. Node-Protection Condition (NPC)

We consider the failure of the node P in Fig. (2). When traffic is rerouted to neighbor N_1, the next hop is again P, the traffic is rerouted to S, and a routing loop occurs. Therefore, N_1 cannot be used as LFA by S to protect the failure of node P. However, N_2 can be used for that objective. A neighbor node N must meet the following node-protection condition (NPC) to protect the failure of a node P:

$$\text{dist}(N, D) < \text{dist}(N, P) + \text{dist}(P, D) \quad (2)$$

An LFA meeting the LFC only is called link-protecting while an LFA also meeting the NPC is called node-protecting. Since the NPC implies the LFC[1], every node-protecting LFA is also link-protecting, but not vice-versa.

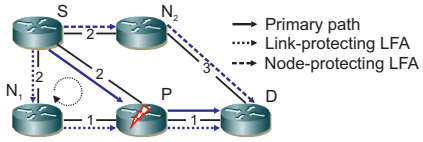


Fig. 2: Only the node-protecting LFA N_2 can be used to protect the failure of node P.

D. Downstream Condition (DSC)

We consider source S and destination D in Fig. (3). N provides a node-protecting LFA for S. If two nodes P_S and P_N fail simultaneously, S reroutes its traffic to N. N cannot forward the traffic, either, and reroutes the traffic to S which is a node-protecting LFA for N in that case. Thus, a routing loop occurs. Such loops which are due to multi-failures can be avoided if an LFA obeys the downstream condition (DSC):

$$\text{dist}(N, D) < \text{dist}(S, D) \quad (3)$$

An LFA fulfilling this condition is called downstream LFA. Allowing only downstream LFAs guarantees loop avoidance for all possible failures because packets get always closer to the destination. In this case, N can be used as LFA for S in Fig. (3) but not vice-versa which avoids the routing loop in our example. N must use another neighbor – if available – to protect against the failure of P_N.

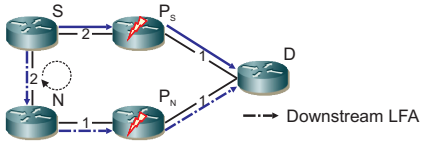


Fig. 3: Neighbor N is a downstream LFA of S but not vice-versa. The use of only downstream LFA avoids loops in the presence of multiple failures.

[1] $\text{dist}(N, D) <^{\text{NPC}} \text{dist}(N, P) + \text{dist}(P, D) \leq^{(a)} \text{dist}(N, S) + \text{dist}(S, P) + \text{dist}(P, D) =^{(b)} \text{dist}(N, S) + \text{dist}(S, D)$ – (a) follows from the triangular equation, (b) holds since the shortest path from S to D leads via P.

E. Equal-Cost Alternates (ECAs)

A special case of LFAs are equal-cost alternates (ECAs), i.e., alternative next hops such that the alternative path is not longer than the primary path. An example is depicted in Fig. (4). The source S knows several paths of equal cost towards D. If its next hop P fails, it can use any of the remaining equal-cost paths as LFA that do not contain the failed element. Thus, either N_1 or N_2 may be used as ECA and even both may be used at the same time. In particular, if the standard routing uses the equal-cost multipath (ECMP) option, the traffic hit by the failure is equally redistributed over the remaining paths.

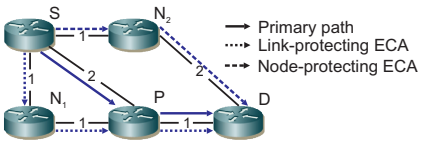


Fig. 4: The equal-cost alternates (ECAs) N_1 and N_2 provide alternate paths with the same length as the primary path. N_1 is just link-protecting while N_2 is node-protecting.

It is easy to see that ECAs cannot create loops in case of multiple failures as they are always downstream LFAs. Therefore, they are link-protecting but not necessarily node-protecting (see N_1 in Fig. (4)). This also shows that downstream LFAs are not necessarily node-protecting.

F. Taxonomy of LFAs

The above conditions limit the number of neighbor nodes as potential LFAs and create thereby sets of neighbors with different ability to protect failures and to avoid loops.

ECAs are always downstream LFAs (DSC). Downstream LFAs (DSC) are always loop-free (LFC). Some neighbor nodes do not meet any of the corresponding conditions. Thus, the set of ECAs is contained in the set of downstream LFAs which is part of the set of general LFAs which are a subset of all neighbor nodes. This relation is depicted in Fig. (5).

The NPC to guarantee node-protecting LFAs is orthogonal to the other conditions: both neighbor nodes in Fig. (4) are ECAs, but only N_2 is node-protecting. N_1 in Fig. (2) and N in Fig. (3) are both downstream LFAs, but only N is node-protecting. N_2 in Fig. (2) is a non-downstream LFA and node protecting while N in Fig. (1) does not meet any condition. Examples for non-downstream non-node-protecting LFAs can also be constructed.

The Venn diagram in Fig. (5) partitions the set of neighbor nodes into 7 different categories. We order them according to a possible preference for their usage as LFAs (the ultimate preference is the network operator's decision [8]):

1) node-protecting ECAs
2) node-protecting downstream LFAs
3) node-protecting LFAs that do not fulfill the downstream condition
4) ECAs that are just link-protecting

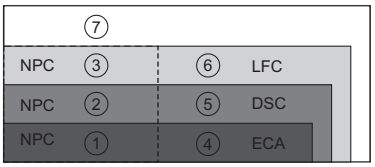


Fig. 5: Classification of neighbor nodes with regard to their ability as forwarding alternates to protect failures and to prevent loops.

5) downstream-LFAs that are just link-protecting
6) LFAs that are just link-protecting and do not fulfill the DSC.

Neighbors not meeting any of the conditions (7) cannot be used as LFAs as they create routing loops.

LFAs cannot achieve 100% failure coverage [14]–[16]. However, they can be complemented by other IP FRR mechanisms with a larger failure coverage.

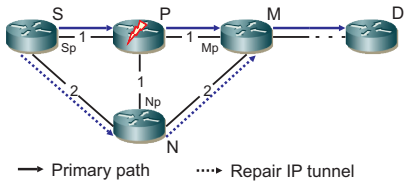
III. IP Fast Reroute Using Not-Via Addresses

The intention of this approach is to protect the failure of a node P or of its adjacent links by deviating affected traffic around P to the next-next hop (NNHOP) M using IP-in-IP tunnelling. The path of this tunnel must not contain the failed node P which is not the case with normal IP forwarding because P is on the shortest path from S to M. Therefore, special "not-via addresses" Mp are introduced such that packets addressed to Mp are forwarded to M not via P. Although the basic idea of IP FRR using not-via addresses is tunnelling to the the NNHOP, it is also possible to protect the last link of a paths with this concept.

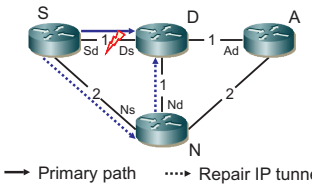
Fig. (6a) illustrates this concept for the case that a NNHOP exists on the primary path. Node S must forward a packet destined to D, but the next hop (NHOP) P (or next link $S \rightarrow P$) fails. Then S encapsulates this packet in another IP packet addressed to the NNHOP using the not-via address Mp. This packet is forwarded from S over N to M which is the shortest path around node P. NNHOP M performs decapsulation and forwards the original packet to D.

Fig. (6b) shows how not-via addresses can be used in case that the NHOP D is already the destination. In contrast to above, node S assumes that only the next link instead of the NHOP has failed; otherwise, the packet cannot be delivered anyway. It encapsulates the packet and addresses it towards Ds. The semantic of Ds at node S is that the direct link $S \rightarrow D$ must not be used. Therefore, the forwarding table at S provides another interface to forward the packet to another neighbor that passes it on to D. Since the packet is sent to Ds, it cannot loop back to S. Finally, D decapsulates the packet and the original packet has reached its destination. If indeed not only link $S \rightarrow D$ but node D has failed, the packet is discarded as soon as it reaches another neighbor of D.

IP FRR using not-via addresses guarantees 100% failure coverage for single node and link failures unless there is an articulation point in the network that splits the network into



(a) An NNHOP exists: encapsulation with address Mp; the encapsulated packet is carried to M not via P.



(b) Next hop is destination: encapsulation with address Ds; the encapsulated packet is forwarded to one of its neighbors and then carried to D not via S, which avoids the use of the failed link $D \rightarrow S$.

Fig. 6: Use of not-via addresses to protect the failure of intermediate nodes and links, and the last link.

two disconnected parts. The concept is very similar to the MPLS FRR facility backup option installing local bypasses to every NNHOP [17]. However, the backup paths in MPLS may follow explicit routes, therefore, MPLS-FRR has more degrees of freedom than IP-FRR using not-via addresses.

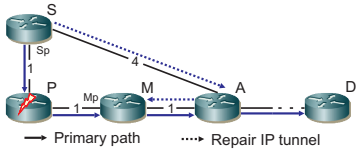


Fig. 7: Unnecessarily long backup paths occur if the bypass from S to the NNHOP M intersects the downstream paths from M to D.

In the example of Fig. (7), packets are carried from S to D over P, M, and A. If P fails, these packets are tunnelled to Mp such that they take the path S, A, M, A, D which is unnecessarily long and wastes capacity, but does not create a loop.

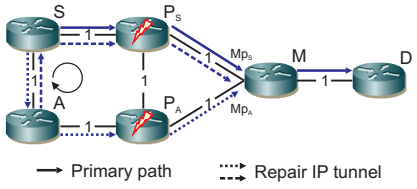


Fig. 8: Routing loops can occur if packets are recursively tunnelled to not-via addresses in case of multiple failures. Hence, recursive tunnelling to not-via addresses is prohibited.

In Fig. (8), S cannot deliver packets to D if nodes P_S and P_A fail. In that case, S encapsulates packets to D in packets destined to Mp_S and these packets are carried to A. A cannot forward the packets to M because P_A also fails. If A encapsulates them to not-via address Mp_A and returns them to S, a routing loop occurs. Therefore, recursive tunnelling using not-via addresses is prohibited [9].

IP FRR using not-via addresses requires the network to provide additional entries in the forwarding tables for not-via addresses. Not-via addresses have the form Mp where p can be any node and M can be any of its neighbors. Therefore, the number of not-via addresses equals the number of unidirectional links in the network. The forwarding entries for the not-via addresses can be constructed by distributed routing algorithms [9].

IV. Comparison of LFAs, Not-Via Addresses, and their Combined Usage

In this section, we compare the pros and cons for LFAs and not-via addresses and discuss how both approaches may be combined.

A. Pros and Cons of LFAs and Not-Via Addresses

In the following, we discuss pros and cons for both approaches.

1) Tunneling: Not-via addresses fully rely on IP tunneling. This involves en- and decapsulation of tunneled packets and may have a performance impact on router hardware. Further, it leads to increased packet lengths inside the tunnel and may result in packet fragmentation due to maximum transmission unit (MTU) limitations. Encapsulation applies a different rewrite string to the front of the packet and most current hardware achieves this without performance degradation. Packet decapsulation at the tunnel endpoint, however, requires two lookup operations. The first to recognize the tunnel endpoint, the second for further forwarding with the inner IP address. Most modern hardware is designed to perform this at line rate. On legacy hardware this can slow down the handling of this specific packet to almost half line rate depending on the router load. So the major disadvantage caused by tunneling stems from packet decapsulation on legacy hardware.

2) Backup Path Length: Since LFAs are computed per destination prefix, they may allow slightly shorter repair paths. While LFAs deviate the packets directly to the destination, not-via addresses deviate the traffic around the failure back onto the original path.

3) Routing Table Size: Not-via addresses require the network to provide additional entries in the forwarding tables. The number of not-via addresses equals the number of unidirectional links in the network. This increase in routing table entries, however, is low compared to the number of entries already present. Some of the entries for not-via addresses are actually unnecessary, since packets destined to not-via addresses will only be seen along the shortest path around the outage location. However, there is no easy way for a router to find out whether it lies on the shortest path for a specific address. LFAs do not require additional entries in the routing table, but each entry for existing destinations must be enhanced with information about the alternate next hop.

4) Computational Routing Complexity: In principle, each node must remove every other node P one by one from the base topology and perform a shortest path tree (SPT) computation in this reduced topology to the not-via addresses Np of P's neighbors N. Incremental SPT (iSPT) computations reduce this effort that is proportional to the number of nodes in the network to an equivalent of 5 to 13 SPT computations in real world networks with 40 to 400 nodes [9]. ECAs in particular are very easy to compute since they are obtained for free from the usual shortest path calculations. For general LFAs, the computational cost of determining individual repair paths for all destinations can be very high as well. So the computational routing complexity and its assessment is hardware- and implementation-dependent.

5) Failure Coverage: If there are no articulation points that disconnect the network in case of a failure, not-via addresses always achieve 100% failure coverage using a single resilience concept. This is usually impossible for LFAs [14]–[16].

6) Compatibility with Loop-Free Re-Convergence Schemes: The computation of the not-via tunnels can be temporally decoupled from the computation of the basic routing. Thus, during routing re-convergence, the tunnels remain stable making not-via addresses compatible with additional mechanisms for loop-free re-convergence [13], [18]. This does not necessarily hold for LFAs since the re-convergence process may render LFA conditions invalid.

7) Protection of Multicast Traffic: Not-via addresses deviate the traffic to the NNHOP through tunnels. Thus, the NNHOP can infer the usual interface from the not-via address and run the reverse path forwarding (RPF) check required for multicast traffic correctly [9]. Protection of multicast traffic with LFAs seems complex and is currently not discussed.

8) Adaptability to SRLGs: The functionality of not-via addresses can be easily adapted to SRLGs. If SRLGs are known, the SPT computation for the respective not-via address is simply performed in the topology with all elements from the SRLG removed. This is much more complicated for LFAs.

Adaptability to SLRGs, protection of multicast traffic, compatibility with loop-free re-convergence schemes and 100% failure coverage are strong advantages in favor of not-via addresses. Possibly shorter backup paths and above all tunneling may have a performance impact and favor the combined usage of LFAs and not-via addresses. In the following we provide further insights into this discussion to assess this tradeoff.

B. Combined Usage of LFAs and Not-Via Routing for Different Resilience Requirements

In this paper, we study three options with different level of failure protection and loop avoidance:

(i) Protection against single link failures
(ii) Protection against single link and single router failures
(iii) Protection against single link and single router failures with loop avoidance in the presence of multiple failures

Not-via addresses fulfill the strictest resilience requirement (iii). LFAs alone cannot even meet the loosest one because they cannot achieve 100% failure coverage, therefore, we complement them by not-via rerouting where necessary. As LFAs have different properties (cf. Fig. (5)), only certain LFA types can be used in the above cases in the following order of preference:

(i) (1), (4), (2), (5), (3), (6), and not-via.
(ii) (1), (2), (3), and not-via; (4), (5), and not-via to protect the last link.
(iii) (1), (2), and not-via; (4), (5), and not-via to protect the last link.

We prefer ECAs over downstream LFAs and downstream LFAs over node- and link-protecting LFAs. For (i) we prefer all LFAs that are node-protecting over link-protecting LFAs. Note that for the protection of the last link for (ii) and (iii) just link-protecting LFAs (6) cannot be used since they may create loops in case the destination node is down.

V. Analysis of the Combined Usage of LFAs and Not-Via Addresses

For the above resilience requirements, we analyze the combined applicability of LFAs and not-via addresses, the backup path prolongation, the amount of decapsulated traffic, the impact on forwarding tables, and the resource requirements in an experimental environment.

A. Experimental Environment

We use well-known realistic networks for our experimental environment: COST239, GEANT, Labnet03, and NOBEL. For compactness sake we only present the results from COST239 (see Fig. (9a)) and from GEANT (see Fig. (9b)) here, since the other networks do not yield additional insights. Those two networks are typical representatives of two different network types. For Labnet03 and Nobel there were quantitative, but no qualitative differences.

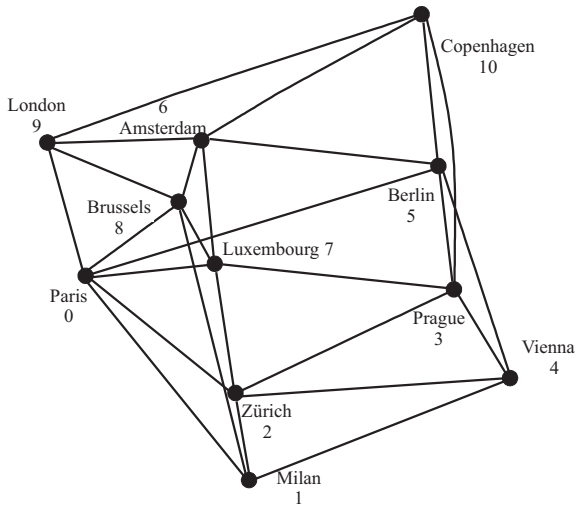
Even for real networks, traffic matrices are generally unavailable due to confidentiality reasons. Thus, we use the method proposed in [19] and enhanced in [20] to generate synthetic traffic matrices resembling real-world data. Note that traffic matrix traces are indeed available for the GEANT network, but we used the synthetically generated traffic matrices here as well to assure comparability.

We set all link weights to one and perform simple hop count routing as often used in unoptimized networks. We perform single shortest path first (SPF) routing. When multiple equal cost paths (ECMPs) towards a destination are available, the interface with the lowest ID is installed as the active interface as specified for IS-IS [21].

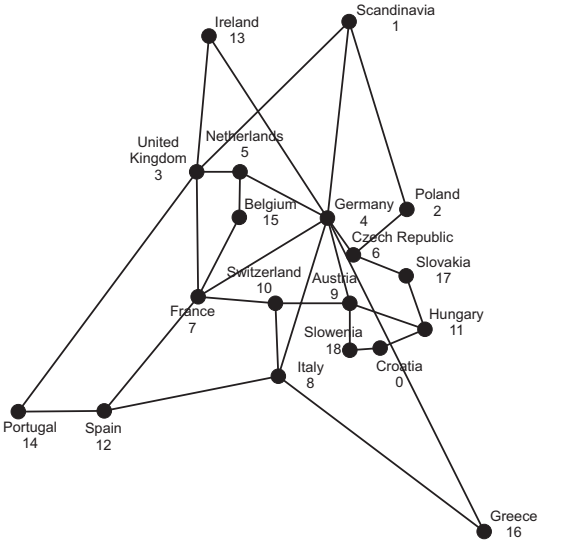
We scaled the traffic matrices such that the maximum link utilization does not exceed 100% for SPF re-convergence and any of the considered failure scenarios.

B. Applicability of LFAs and Not-Vias

We first study the applicability of LFAs and not-vias at the individual network nodes. Fig. (10) shows the percentage



(a) COST239



(b) GEANT

Fig. 9: Networks under study.

of the destinations protected by different types of LFAs and not-vias for the 11 nodes in the COST239 network and resilience requirements (i) to (iii). Fig. (11) contains the same information for the 19 nodes of the GEANT network. The x-axes show the node IDs and the y-axes the percentage of destinations at a node covered by the respective mechanism in percent. We applied appropriate LFAs and not-via protection according to the recommendations in Section IV-B. Since there is a slightly different semantic (cf. Section III) for not-via addresses for the last hop, we indicate not-vias used for the protection of the last hop towards a destination separately.

We first start with general observations. In networks using simple hop count routing, only three out of six types of neighbors (cf. Fig. (5)) providing LFAs exist. First, ECAs that are only link-protecting (4) do not exist since there are no parallel links. Second, there are no downstream LFAs (2),(5). The downstream criterion requires that the alternate neighbor N is closer to the destination D than the deviating node S. Since the distance dist(S, N) from S to its neighbor N is always 1, this can only be true for equal cost paths.

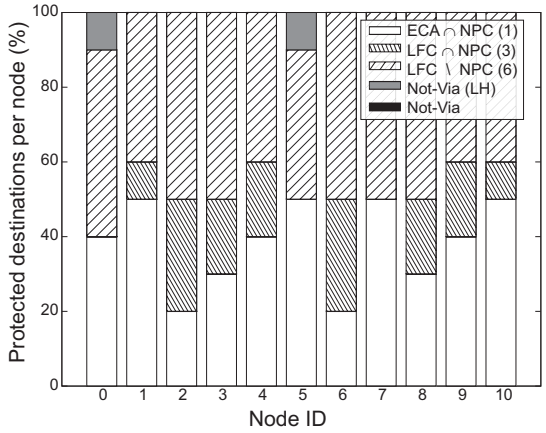
We now discuss the results from the COST239 network. The COST239 topology represents a class of networks that are well connected among the individual nodes. For most nodes any other node is reachable within at most two hops. In Fig. (10a) corresponding to resilience requirement (i) – link protection only – almost all destinations can be protected using LFAs. ECAs (1) protect between 20-50% of the destinations and node-protecting LFAs (3) vary from 0 to 30%. Link-protecting LFAs (6) are applicable for a high percentage of destinations between 40 - 50 %, mainly to protect the last hops of the relatively short paths. Almost no not-vias are necessary. Only two nodes require about 10% of not-vias for the last hop.

Fig. (10b) shows the results for the stricter resilience requirement (ii) – link and node protection. All link-protecting LFAs (6) are replaced with not-vias. For the strictest resilience requirement (iii) – link and node protection with general loop avoidance – shown in Fig. (10c), node-protecting LFAs (3) are not sufficient anymore and are again replaced by not-vias. Now, only ECAs and not-vias are applicable due to the non-existence of downstream LFAs.

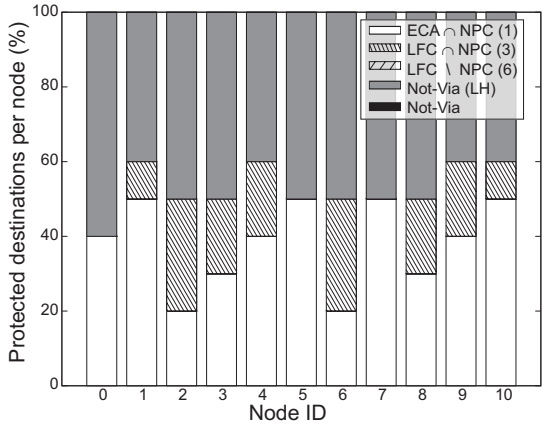
The GEANT topology in contrast represents a more sparsely connected class of network topologies. The paths between node pairs are significantly longer since the nodes lie on circles of three to five nodes. Concerning the results, the variation between the individual nodes is high. In Fig. (11a) for resilience requirement (i), node 16 is very different from the other nodes. It uses 100% link-protecting LFAs (6). This can be explained by its special location forming a triangle with nodes 4 and 8. Besides node 16, only two other nodes use these LFAs (6) while the number of node-protecting LFAs (3) varies greatly between 0 and almost 80%. In contrast to the COST239 network, all nodes except for node 16, require not-vias for the protection of the last hops, and up to 70% of all destinations within a node's routing table can only be protected using not-via addresses.

For resilience requirement (ii) in Fig. (11b), again all link-protecting LFAs (6) cannot be used anymore. Consequently, node 16 requires 100% not-vias. For the strictest resilience requirement (iii) in Fig. (11c), again only ECAs (1) and not-vias are applicable. Now node 16 and 17 require 100% not-vias.

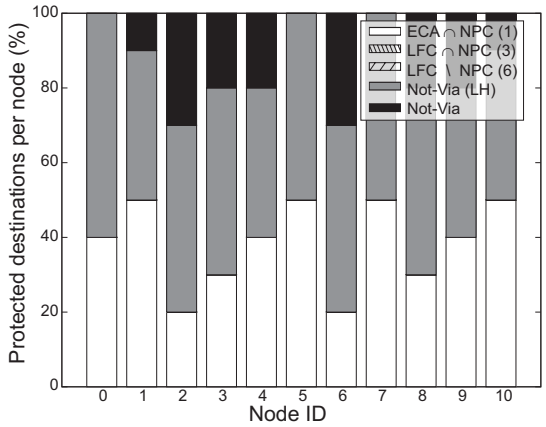
The conclusion from this analysis is threefold. (1) In case of simple hop count routing three out of six types of LFAs do not exist. (2) If loop avoidance in general failure cases is required (iii), LFAs other than ECAs cannot be used in networks that use simple hop count routing. (3) Average values for the coverage achieved by LFAs as shown in previous work is not a sufficient performance metric: the existence of suitable LFAs largely depends on the network topology and in certain topologies individual nodes cannot protect a single destination under resilience requirements (ii) and (iii) with LFAs. The average values hide these variations. Hence, not-vias are not only necessary as an additional FRR mechanism for LFAs for



(a) Link protection only.

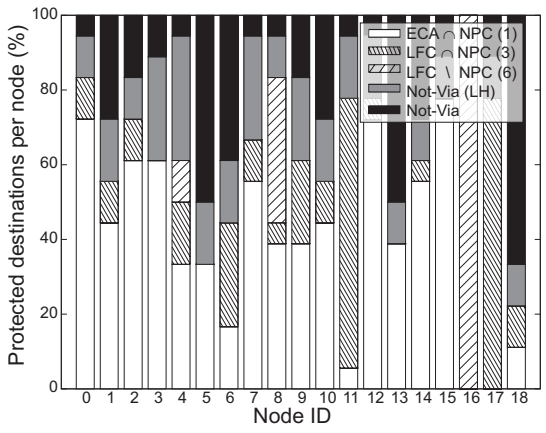


(b) Link and node protection.

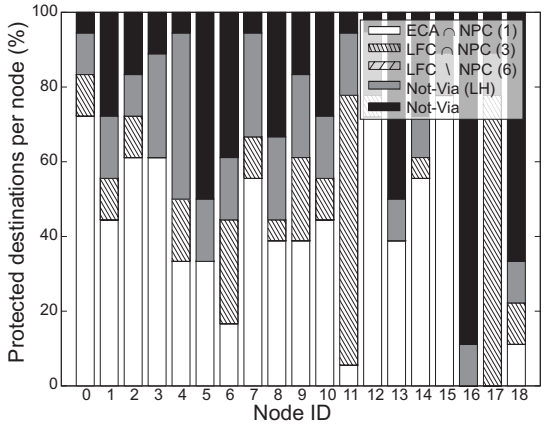


(c) Link and node protection - no loops during multiple failures.

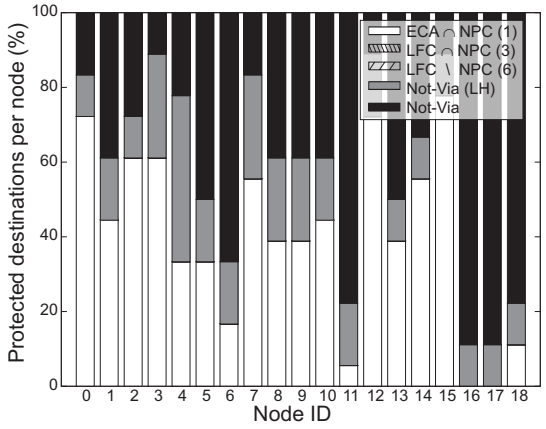
Fig. 10: Applicability of LFAs and not-via addresses in the COST239 network with different resilience requirements.



(a) Link protection only.



(b) Link and node protection.



(c) Link and node protection - no loops during multiple failures.

Fig. 11: Applicability of LFAs and not-via addresses in the GEANT network with different resilience requirements.

100% coverage, at some nodes they are the only option.

C. Path Prolongation

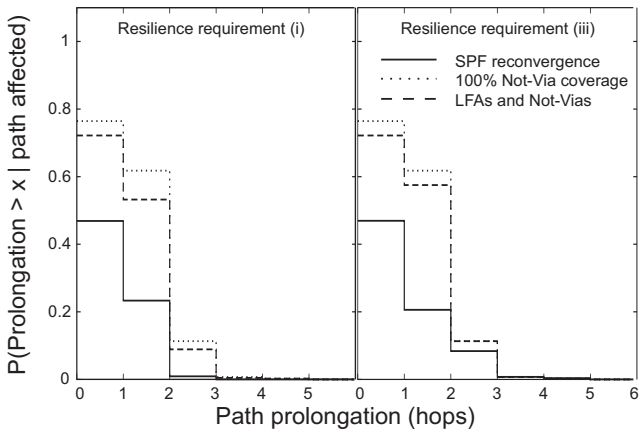


Fig. 12: Path prolongation in the GEANT network for resilience requirements (i) and (iii).

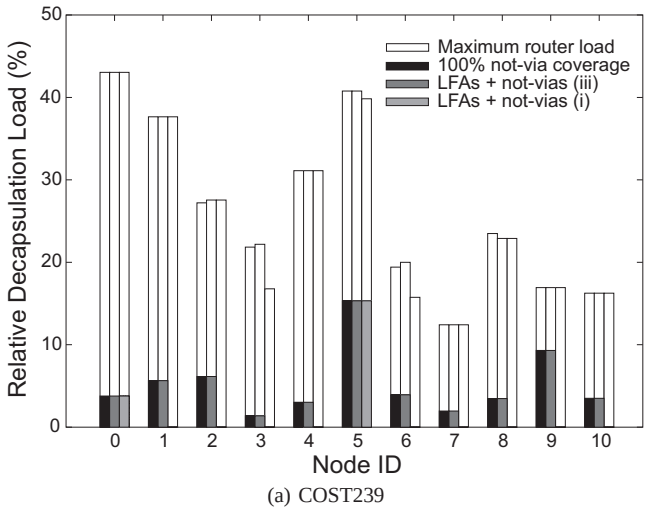
The backup path should not be much longer than the original path for delay sensitive applications. Hence, we assess the path prolongation for all failure scenarios. Fig. (12) shows the CCDF for the path prolongation for resilience requirement (i) and (iii) in the GEANT network. The x-axes shows the path prolongation x in number of hops, the y-axes shows the conditional probability that a path affected by a failure increases by more than x hops. SPF re-convergence is the comparison baseline since the backup path cannot be shorter.

The length of about 50% of the paths does not increase for plain IP re-convergence. These are the paths where alternative paths of equal length exist between source and destination. This value decreases to around 25% if IP FRR is applied since fewer ECAs are available for local repair at intermediate nodes. The difference between IP FRR and SPF re-convergence is well noticable, however, the difference between 100% not-via coverage and the combination of LFAs with not-vias is small and well tolerable. This difference even decreases for the strictest resilience requirement (iii).

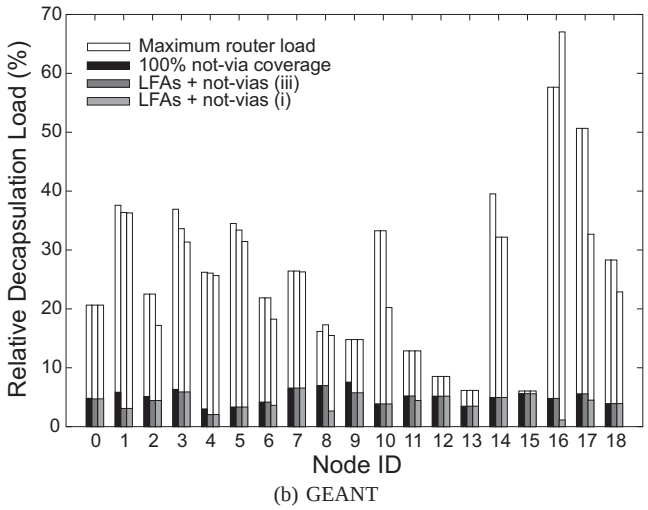
We omit the values for the COST239 network since there is no difference between both IP FRR mechanisms, the difference between SPF and IP FRR is similar to GEANT.

D. Decapsulated Traffic from Not-Via Tunnels

In Figs. (13a)–(13b) we analyze the amount of traffic that must be decapsulated at the not-via tunnel endpoints. All numbers for the individual nodes are relative to the node capacity, which is the sum of the capacity of the incoming interfaces of the node. Our performance metric of interest is the maximum amount of decapsulated traffic observed in all protected failure scenarios. The bars in the background show the maximum amount of incoming traffic, i.e., the maximum router load, to relate the results to the overall traffic at a node. Note that the maximum router load is well below 100% since the load reaches its maximum for individual incoming links in different scenarios.



(a) COST239



(b) GEANT

Fig. 13: Amount of decapsulated traffic per node relative to maximum node capacity for COST239 and GEANT.

In COST239 (Fig. (13a)) with 100% not-via coverage, almost all nodes must decapsulate at most traffic well below 10% of their capacity. Only node 5 shows a higher value of 15%. Surprisingly, there is no reduction of the maximum amount of decapsulated traffic with the combined usage for resilience requirement (iii). This does not mean that the deployment of LFAs does generally not reduce the amount of decapsulated traffic in all failure scenarios, but the maximum amount cannot be reduced here. For combined coverage and resilience requirement (i), only nodes 0 and 5 still decapsulate packets. These are the only two nodes that require not-vias to protect 100% of their destinations. Interestingly, node 0 tunnels packets to node 5 and vice versa. This phenenomen is due to the network structure. While all other pairs of neighboring nodes form triangles with a third node allowing to use a link-protecting LFA, for nodes 0 and 5 only a quadrangle can be found. Again, the maximum amount of decapsulated traffic does not decrease at those two nodes.

The results are slightly different in the GEANT network (Fig. (13b)). The maximum values stay well below 8% of the node capacities. For combined usage and resilience requirement (iii), the maximum amount of decapsulated traffic reduces for one half of the nodes, but most nodes show only small differences. For resilience requirement (i) a further reduction is noticable for individual nodes, especially nodes 8 and 16, but all nodes must still decapsulate traffic.

In general, the combined usage of LFAs and not-vias does not reduce the maximum amount of decapsulated traffic much. In particular, if more than pure link protection is required.

E. Impact on Routing Tables

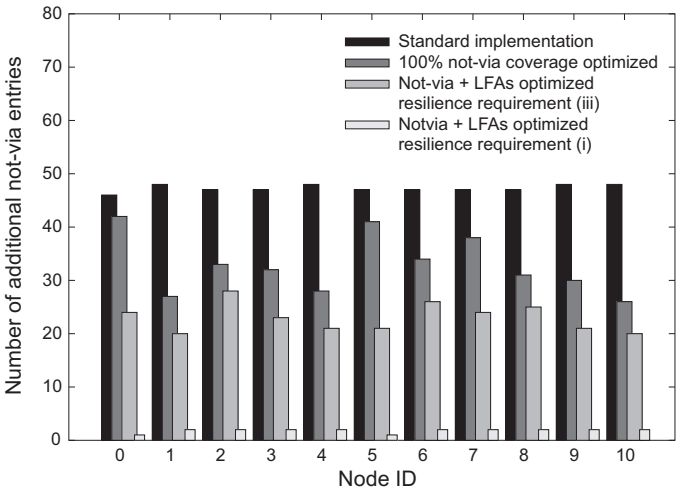
IP FRR using not-vias requires the network to provide additional entries in the forwarding tables for not-via addresses. Therefore, we assess the impact of these additional entries in Figs. (14a)–(14b) for the COST239 and the GEANT network. The x-axes correspond to the node IDs, the y-axes show the actual number of additional entries required in the forwarding tables of the individual nodes due to not-via addresses.

Before we discuss the results from the graphs in detail, we start with a few general observations. The number of additional not-via addresses in the networks equals the number of unidirectional links (cf. Section III). In theory, not all nodes need to add the entire set of not-via addresses to their forwarding tables. Only the nodes along a not-via repair tunnel must know the corresponding not-via address. However, there is no simple way to detect whether a node lies along a repair tunnel, i.e., whether it is on a shortest path for a specific not-via address. Further, LFAs make a not-via tunnel obsolete if and only if all traffic sent through this tunnel is protected by LFAs instead. Since LFAs are locally computed per destination prefix, there is no simple way for an arbitrary node to detect this. An arbitrary node must check all destinations that are protected by the considered not-via tunnel in the forwarding table of the tunnel starting point whether they can be protected by LFAs instead. This is clearly complex.

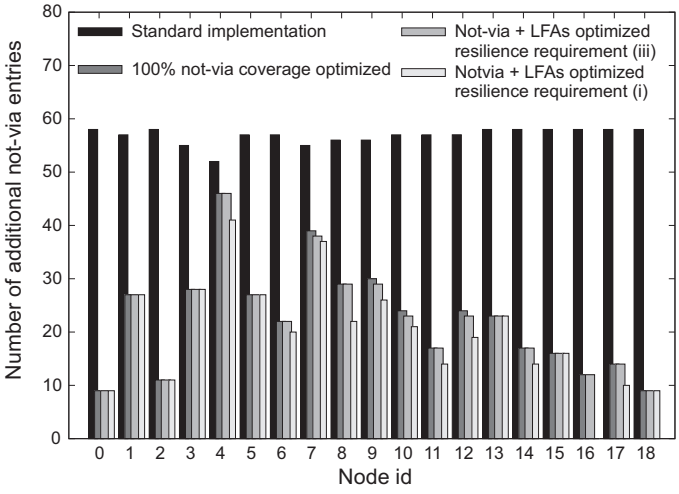
Due to these complexity considerations it is unlikely that an implementation of not-via addresses and LFAs in practice checks whether a not-via address entry is required in a forwarding table or not. Therefore, out straightforward standard implementation simply creates one entry for each not-via address in all forwarding tables for both 100% not-via protection and not-vias in combination with LFAs. No entry is required for the not-via addresses advertised by a node itself corresponding to the number of incoming links.

The black bars in the graphs represent this standard implementation. Note that node 0 in the COST239 network and node 4 in the GEANT network require the least additional entries since they have the largest number of incoming links.

Still, there is a theoretical optimization potential. Thus, to assess the actually required number of additional entries and the potential of LFAs to further reduce it, we used a simple brute force method. We check all repair paths in all considered single failure scenarios node by node to see which additional entries are required.



(a) COST239



(b) GEANT

Fig. 14: Number of additional entries required in the forwarding tables of individual nodes for the COST239 and the GEANT network.

Fig. (14a) reveals the following results for COST239. With 100% not-via coverage, between 8% to 50% of the not-via entries of a node may be deleted from the routing tables since they will never be used. The combination of not-via addresses and LFAs for resilience requirement (iii) further reduces the required entries for most nodes only slightly. The combination of both mechanisms for the weakest resilience requirement (i) reduces the amount of not-via entries that are absolutely necessary to two. Those are the addresses "0 not-via 5" and "5 not-via 0" for the not-via traffic from 0 to 5 and vice versa if the link between nodes 0 and 5 fails. Hence, nodes 0 and 5 store only one entry.

We briefly go a bit deeper into the last observation. The shortest path between node 0 and node 5 goes over a single link, but when it fails, the shortest path now requires 3 hops (cf. Fig. (9a)). However, there are multiple paths of length 3 and in fact, due to the special structure of the COST239

network, each node lies on a shortest path from node 0 to node 5 and vice versa. In case of single shortest path routing, the not-via traffic is forwarded only over one of these equal-cost shortest backup path. However, the choice over which path the not-via traffic is deviated is often random in practice. Therefore, all nodes lying on one of several equal-cost shortest paths require the not-via entries. Unfortunately, these are all nodes of the COST239 network in this case.

The results in Fig. (14b) for the GEANT network are different. Up to 85% of the entries of a node may be deleted from the routing tables since they will never be used for 100% not-via coverage. The difference to COST239 is due to the lower amount of equal cost paths in the network. However, the deployment of LFAs in combination with not-vias hardly leads to a further reduction for both resilience requirements. This is due to the relatively large amount of destinations that can only be protected with not-via addresses even for the weakest resilience requirement (i).

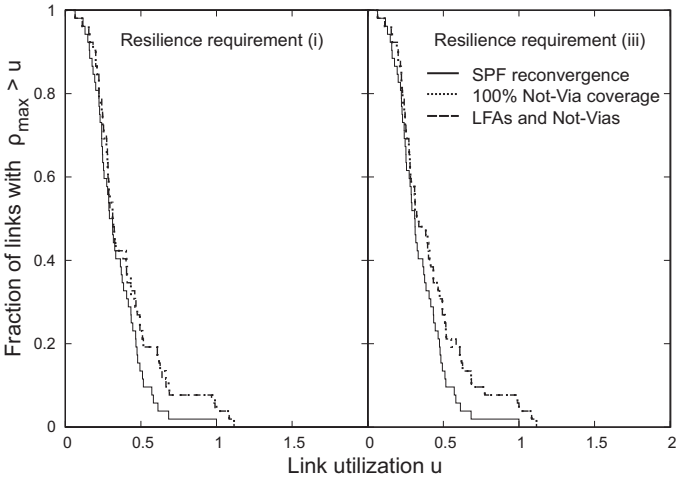
So the result of the analysis of the number of additional entries in the forwarding tables is twofold. (1) There is only a slight impact caused by not-via addresses. One entry per unidirectional link in the network is not a large number compared to the much larger number of prefixes usually installed in the forwarding tables. (2) The reduction due to the deployment of LFAs is computationally complex. Besides, it is noticeable only in networks where all nodes can protect a large amount of destinations using LFAs.

F. Maximum Link Utilization

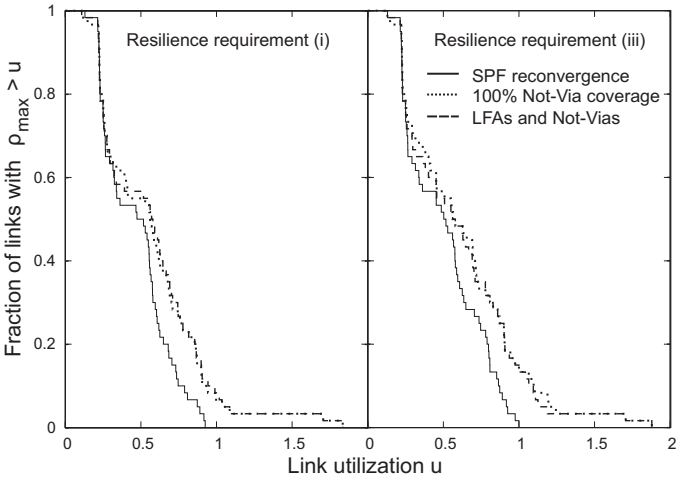
Finally, we analyze the resource requirements of the IP FRR concepts under study. To that purpose we calculated for each link in the network the maximum link utilization over the considered failure scenarios, i.e., over all single link failures for resilience requirement (i) and all single link and single node failures for resilience requirement (iii). The results are shown in Fig. (15a) for the COST239 and in Fig. (15b) for the GEANT network. The x-axes show the link utilization u, the y-axes show the fraction of the links in the network with maximum link utilization $\rho_{max} > u$. The graphs show the results for SPF re-convergence as comparison baseline and 100% not-via coverage and not-vias in combination with suitable LFAs. The numbers are normalized such that the maximum utilization value for SPF re-convergence and requirement (iii) reaches 100%.

On most links, the deployment of IP FRR mechanisms increases the maximum link utilization in both networks by at most 0.2 relative to SPF re-convergence. Only a view links suffer from large additional capacity requirements. Interestingly, there is virtually no difference between 100% not-via coverage and the combined application independent of the resilience requirement. Specifically, for COST239 there is no difference visible, for GEANT there is only a slight difference.

Concerning the difference between resilience requirement (i) and (iii), the networks behave differently. In the COST239 network (cf. Fig. (15a)), the graphs are very similar for both resilience requirements. The protection against single link failures (i) as shown in the left half of the figure requires only slightly less capacity than protection against single link and node failures (iii) as shown in the right half. This is due to the short paths in this network. In the GEANT network (cf. Fig. (15b)), link failure protection (i) requires less capacity than link and node failure protection (iii), but the qualitative difference between the mechanisms under study does not change.



(a) COST239



(b) GEANT

Fig. 15: Fraction of links with maximum link utilization ρ_{max} greater than a given value u for different resilience requirements and resilience mechanisms in the COST239 and the GEANT network.

Overall, there is a price to pay in terms of maximum link utilization for the deployment of IP FRR mechanisms. However, a significant difference between 100% not-via coverage and the combined application with LFAs cannot be found.

VI. Related Work

Fast reroute (FRR) concepts were first developed for MPLS technology and standardized in [6]. Currently, extensions for point-to-multipoint are under discussion to protect multicast

traffic [22], [23]. For IP routing, its ability for sub-second reaction by adjusting timer values and stability issues when performing such optimizations were studied in, e.g., [5], [24]. [7] provides a framework for IP FRR currently under development by the IETF routing working group (RTGWG). This group also published Internet drafts proposing LFAs [8] and not-via addresses [9]. Among other concepts for IP FRR are multiple routing configurations (MRC) and failure inferencing based fast rerouting.

Multiple routing configurations (MRC) described in [10] and as a similar concept in [25], [26] are a small set of backup routing configurations for use in failure cases. The routing configurations 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 configuration. This concept can be implemented using the multi-topology extensions for OSPF and IS-IS [27]–[29]. [30] proposed an extension called 2DMRC to handle concurrent multi-failures with MRC.

Failure inferencing based fast rerouting (FIFR) exploits the fact that packets arrive at routers through other interfaces during network element failures if rerouting is applied than in case of normal operation. It computes interface-specific forwarding tables where the next hop of a packet does not only depend on its destination address but also on the incoming interface. It has been proposed to handle transient link [11] and node [31] failures. The original mechanism had problems with assymetric link weights, but this has been fixed in [32] where extensions to handle inter-AS failures have also been developed. [33] suggested a modification called blacklist-based interface-specific forwarding (BISF) that avoids routing loops also in case of multiple failures.

In the context of IP FRR the authors of [34] developed a method to achieve fast recovery of BGP peering link failures. Important are also concepts for loop-free re-convergence that can be used in combination with IP FRR mechanisms in case of long-lived failures. [13] provides a framework for it. One possible suggestion for loop-free reconvergence specifies an order in which nodes are allowed to update their forwarding tables in case of outages and after failure repair or installation of new network elements [35], [36].

VII. Conclusion

In this work we studied the combined usage of two IP FRR mechanisms currently under standardization by the IETF: loop-free alternates (LFAs) and not-via addresses. In case of failures, LFAs deviate traffic to neighboring nodes providing an alternate path towards the destination that avoids the failed element and does not create loops. Not-via addresses bypass the failed element with local IP-in-IP tunnels.

We classified different sets of neighbors providing LFAs according to their ability and established a new taxonomy for LFAs. This taxonomy suggests an order of preferred combinations of LFAs and not-vias for three types of resilience requirements presented in this paper.

LFAs alone cannot achieve 100% failure coverage and must be complemented by other IP-FRR mechanisms like not-vias. Our analysis of their combined usage revealed that three out of six types of LFAs do not exist in networks using simple hop count routing. If single link and node failures should be protected, at least 50% of all destinations of a node require not-via protection on average. Depending on the network topology, the variation between indivudal nodes can be very high, leading to nodes that cannot protect a single destination without not-via addresses.

IP FRR mechanisms lead to longer backup paths than plain IP re-convergence. However, the combined usage of LFAs and not-via addresses leads only to slightly shorter backup paths than 100% not-via coverage. The same holds for the maximum amount of decapsulated traffic caused by not-via tunneling. The combined usage cannot reduce this amount significantly.

The reduction of required additional entires for not-via addresses in the forwarding tables when combining LFAs with not-vias is computationally complex and only noticeable in networks where all nodes can protect a large amount of destinations using LFAs.

Finally, there is a price to pay in terms of resource requirements for the deployment of IP FRR mechanisms relative to plain IP re-convergence, but there is no difference between 100% not-via protection and combined deployment.

These findings support the following recommendation. If 100% failure coverage with IP FRR is required, not-via addresses should be applied as the only FRR mechanism since our results show no strong advantages of the combined application. A homogeneous solution also leads to a simpler network management.

References

[1] G. Iannaccone, C.-N. Chuah, R. Mortier, S. Bhattacharyya, and C. Diot, "Analysis of Link Failures in an IP backbone," in *ACM SIGCOMM Internet Measurement Conference*, Marseille, France, Nov. 2002, pp. 237–242.

[2] A. Nucci, B. Schroeder, S. Bhattacharyya, N. Taft, and C. Diot, "IGP Link Weight Assignment for Transient Link Failures," in 18^{th} *International Teletraffic Congress (ITC)*, Berlin, Germany, Sep. 2003.

[3] B. Fortz and M. Thorup, "Robust Optimization of OSPF/IS-IS Weights," in *International Network Optimization Conference (INOC)*, Paris, France, Oct. 2003, pp. 225–230.

[4] V. Jacobson, C. Alaettinoglu, and H. Yu, "Towards Milli-Second IGP Convergence," http://tools.ietf.org/id/draft-alaettinoglu-isis-convergence-00.txt, Nov. 2000.

[5] A. Basu and J. G. Riecke, "Stability Issues in OSPF Routing," in *ACM SIGCOMM*, San Diego, CA, USA, Aug. 2001, pp. 225–236.

[6] P. Pan, G. Swallow, and A. Atlas, "Fast Reroute Extensions to RSVP-TE for LSP Tunnels," http://www.ietf.org/rfc/rfc4090.txt, May 2005.

[7] M. Shand and S. Bryant, "IP Fast Reroute Framework," http://www.ietf.org/internet-drafts/draft-ietf-rtgwg-ipfrr-framework-07.txt, Jun. 2007, expires December 2007.

[8] A. Atlas and A. Zinin, "Basic Specification for IP Fast-Reroute: Loop-free Alternates," http://www.ietf.org/internet-drafts/draft-ietf-rtgwg-ipfrr-spec-base-10.txt, Nov. 2007, expires May 19, 2008.

[9] S. Bryant, M. Shand, and P. S., "IP Fast Reroute Using Not-via Addresses," http://www.ietf.org/internet-drafts/draft-bryant-shand-ipfrr-notvia-addresses-03.txt, Jul. 2007, expires January 2008.

[10] A. Kvalbein, A. F. Hansen, T. Cicic, S. Gjessing, and O. Lysne, "Fast IP Network Recovery using Multiple Routing Configurations," in *IEEE Infocom*, Barcelona, Spain, Apr. 2006.

[11] S. Nelakuditi, S. Lee, Y. Yu, Z.-L. Zhang, and C.-N. Chuah, "Fast Local Rerouting for Handling Transient Link Failures," *IEEE/ACM Transactions on Networking*, vol. 15, no. 2, pp. 359–372, Jun. 2007.
[12] A. Markopoulou, G. Iannaccone, S. Bhattacharyya, C.-N. Chuah, and C. Diot, "Characterization of Failures in an IP Backbone," in *IEEE Infocom*, Hong Kong, Mar. 2004.
[13] S. Bryant and M. Shand, "A Framework for Loop-free Convergence," http://www.ietf.org/internet-drafts/draft-ietf-rtgwg-lf-conv-frmwk-01.txt, Jun. 2007, expires December 2007.
[14] P. Francois and O. Bonaventure, "An Evaluation of IP-based Fast Reroute Techniques," in *CoNEXT (formerly QoFIS, NGC, MIPS)*, Toulouse, France, Oct. 2005, pp. 244–245.
[15] A. F. Hansen, T. Cicic, and S. Gjessing, "Alternative Schemes for Proactive IP Recovery," in 2^{nd} *Conference on Next Generation Internet Design and Engineering (NGI)*, Valencia, Spain, Apr. 2006.
[16] M. Gjoka, V. Ram, and X. Yang, "Evaluation of IP Fast Reroute Proposals," in *IEEE International Conference on COMmunication System softWAre and MiddlewaRE (COMSWARE)*, Bangalore, India, Jan. 2007.
[17] R. Martin, M. Menth, and K. Canbolat, "Capacity Requirements for the Facility Backup Option in MPLS Fast Reroute," in *IEEE Workshop on High Performance Switching and Routing (HPSR)*, Poznan, Poland, Jun. 2006.
[18] P. Francois, M. Shand, and O. Bonaventure, "Disruption-Free Topology Reconfiguration in OSPF Networks," in *IEEE Infocom*, Anchorage, Alaska, USA, May 2007.
[19] A. Nucci, A. Sridharan, and N. Taft, "The Problem of Synthetically Generating IP Traffic Matrices: Initial Recommendations," *ACM SIGCOMM Computer Communications Review*, vol. 35, no. 3, pp. 19–32, Jul. 2005.
[20] M. Roughan, "Simplifying the Synthesis of Internet Traffic Matrices," *ACM SIGCOMM Computer Communications Review*, vol. 35, no. 5, pp. 93–96, Oct. 2005.
[21] Digital Equipment Corp., "RFC1142: OSI IS-IS Intra-domain Routing Protocol," ftp://ftp.rfc-editor.org/in-notes/rfc1142.pdf, Feb. 1990.
[22] R. Aggarwal, D. Papadimitriou, and S. Yasukawa, "Extensions to Resource Reservation Protocol - Traffic Engineering (RSVP-TE) for Point-to-Multipoint TE Label Switched Paths (LSPs)," http://www.rfc-editor.org/rfc/rfc4875.txt, May 2007.
[23] J. L. Le Roux, R. Aggarwal, J. P. Vasseur, and M. Vigoureux, "P2MP MPLS-TE Fast Reroute with P2MP Bypass Tunnels," http://www.ietf.org/internet-drafts/draft-ietf-mpls-p2mp-te-bypass-01.txt, Jul. 2007, expires: January 2007.
[24] P. Francois, C. Filsfils, J. Evans, and O. Bonaventure, "Achieving Sub-Second IGP Convergence in Large IP Networks," *ACM SIGCOMM Computer Communications Review*, vol. 35, no. 2, pp. 35–44, Jul. 2005.
[25] G. Apostolopoulos, "Using Multiple Topologies for IP-only Protection Against Network Failures: A Routing Performance Perspective," Institute of Computer Science (ICS) of the Foundation for Research and Technology - Hellas (FORTH), Heraklion, Crete, Greece, Tech. Rep. TR377, 2006.
[26] M. Menth and R. Martin, "Network Resilience through Multi-Topology Routing," in 5^{th} *International Workshop on Design of Reliable Communication Networks (DRCN)*, Island of Ischia (Naples), Italy, Oct. 2005, p. 271 90 277.
[27] P. Psenak, S. Mirtorabi, A. Roy, N. L., and P.-E. P., "Multi-Topology (MT) Routing in OSPF," http://www.ietf.org/rfc/rfc4915.txt?number=4915, Jun. 2007.
[28] N. Rawat, R. Shrivastava, and D. Kushi, "OSPF Version 2 MIB for Multi-Topology (MT) Routing," ftp://ftp.rfc-editor.org/in-notes/internet-drafts/draft-ietf-ospf-mt-mib-01.txt, Aug. 2007, expires February 23, 2008.
[29] T. Przygienda and N. Shen, "Multi Topology (MT) Routing in IS-IS," http://www.ietf.org/internet-drafts/draft-ietf-isis-wg-multi-topology-12.txt, Nov. 2007, expires: May 2008.
[30] A. F. Hansen, O. Lysne, T. Cicic, and S. Gjessing, "Fast Proactive Recovery from Concurrent Failures," in *IEEE International Conference on Communications (ICC)*, Glasgow, UK, Jun. 2007.
[31] Z. Zhong, S. Nelakuditi, Y. Yu, S. Lee, J. Wang, and C.-N. Chuah, "Failure Inferencing based Fast Rerouting for Handling Transient Link and Node Failures," in *IEEE Global Internet Symposium*, Miami, Fl, USA, Mar. 2005, pp. 2859–2863.
[32] J. Wang and S. Nelakuditi, "IP Fast Reroute with Failure Inferencing," in *ACM SIGCOMM Workshop on Internet Network Management (INM)*, Kyoto, Japan, Aug. 2007.
[33] J. Wang, Z. Zhong, and S. Nelakuditi, "Handling Multiple Network Failures through Interface Specific Forwarding," in *IEEE Globecom*, San Francisco, CA, USA, Nov. 2006.
[34] O. Bonaventure, C. Filsfils, and P. Francois, "Achieving Sub-50 Milliseconds Recovery Upon BGP Peering Link Failures," in *CoNEXT (formerly QoFIS, NGC, MIPS)*, Paris, France, Jul. 2005.
[35] P. Francois, O. Bonaventure, M. Shand, S. Bryant, and S. Previdi, "Loop-Free Convergence Using Order FIB Updates," http://www.ietf.org/internet-drafts/draft-ietf-rtgwg-ordered-fib-01.txt, Jul. 2007, expires January, 2008.
[36] P. Francois and O. Bonaventure, "Avoiding Transient Loops during IGP Convergence in IP Networks," in *IEEE Infocom*, Miami, Florida, Mar. 2005.