Optimizing Post-Failure Network Performance for IP Fast Reroute using Tunnels

Kin-Hon Ho, Ning Wang
University of Surrey
Guildford, Surrey, United Kingdom
K.Ho,N.Wang@surrey.ac.uk

George Pavlou
University College London
London, United Kingdom
G.Pavlou@ee.ucl.ac.uk

Christos Botsiaris
University of Surrey
Guildford, Surrey, United Kingdom
Cbotsiaris@yahoo.gr

ABSTRACT

IP Fast ReRoute (FRR) mechanisms have been proposed to achieve fast failover for supporting Quality of Services (QoS) assurance. However, these mechanisms do not consider network performance after affected traffic is rerouted onto repair paths. As a result, QoS deterioration may still happen due to post-failure traffic congestion in the network, which nullifies the effectiveness of IP FRR. In this paper, by considering IP tunneling as the underlying IP FRR mechanism, we proposed an efficient algorithm to judiciously select tunnel endpoints such that the network performance is optimized after the repair paths are activated for rerouting. According to the simulation results using real operational network topologies and traffic matrices, the algorithm achieves significant improvement on post-failure load balancing compared to the traditional IGP re-convergence and plain tunnel endpoint selection without such consideration.

Categories and Subject Descriptors

C.2.3 [**Computer-Communication Networks**]: Network Operations – *Network Management*

General Terms

Algorithms, Management, Performance, Reliability, Experimentation

Keywords

Fast Reroute, Traffic Engineering

1. INTRODUCTION

Any network Quality of Service (QoS) degradation can lead to perceived service disruption by end customers, which would result in financial and reputation penalties on the offering ISPs. QoS degradation can be attributed to many reasons. One common cause of such deterioration is network failure, which becomes part of daily operations in IP networks [1]. In an IP-based network, when a link or router fails, Interior Gateway Protocol (IGP) re-convergence process takes place to maintain consistent routing. However, it has been shown that network-wide IGP re-convergence may take up to several minutes to complete [3]. During this period, individual routers may have inconsistent views on the overall network topology and therefore transient forwarding loops can be formed [3]. An effective solution has been proposed to recover routing failures in a very short time to avoid noticeable service disruptions: Once a router detects the failure of its adjacent network component (e.g. a link or a neighboring router), it immediately reroutes the affected traffic to a pre-computed repair path through which the traffic is forwarded to the destination, with the dissemination of link state advertisement (LSA) for the failure suppressed. This solution is known as IP Fast Re-Route (FRR).

It should be noted that IP FRR does not consider post-failure traffic re-optimization, for instance how to re-balance the overall traffic loading after the affected traffic is re-routed onto the repair paths. Without such consideration on traffic control across individual repair paths, although failures can be bypassed rapidly, there could be an overwhelming amount of traffic re-routed through some repair paths, which leads to congestion on some parts of the network and eventually causes packet delay or loss. This nullifies the effectiveness of IP FRR and hinders the support of QoS assurance. To provide genuine QoS assurance under failures, not only fast routing recovery techniques, but also the provisioning of repair paths that optimizes post-failure network performance should be considered in conjunction.

In this paper we introduce a tunnel-based mechanism as the underlying IP FRR platform. The mechanism makes use of intermediate routers, often known as tunnel endpoints to re-route traffic towards the final destination without traversing the failed network component. To perform FRR after the detection of failure, a router that is adjacent to the failure (called repairing router), tunnels the affected traffic to a tunnel endpoint from where the traffic is de-capsulated and forwarded natively to the final destination. It is worth mentioning that for a given network topology and IGP link weight configuration, multiple intermediate routers may exist as feasible tunnel endpoint candidates to avoid using the failed network component. Hence, an opportunity exists for the network operator to perform optimized selection of tunnel endpoints for achieving post-failure load balancing, provided that the overall traffic matrix can be estimated *a priori*.

We propose an efficient optimization algorithm for the tunnel endpoint selection in order to achieve a comprehensive paradigm for supporting high QoS assurance. The ultimate objective is to minimize the maximum link utilization that takes into account every single link failure scenario. More specifically, based on the overall network topology and the estimated traffic matrix, a tunnel endpoint is selected for each affected destination with regard to each link to be protected. The goal is to re-balance

QShine'08, July 28–31, 2008, Hong Kong, China.
Copyright 2008 ICST ISBN 978-963-9799-26-4
DOI 10.4108/ICST.QSHINE2008.3950

the overall traffic loading after the traffic is rerouted over the repair paths. All the selected tunnel endpoints need to be pre-configured by the network operator at each individual repairing router such that they can be immediately activated once the failure of the protected network component is detected. We evaluate our tunnel endpoint selection algorithm by simulation using real operational network topologies and traffic matrices. The evaluation shows that our algorithm achieves significant improvement on post-failure load balancing compared to the traditional IGP re-convergence and also plain tunnel endpoint selection without such consideration.

The paper is organized as follows. The next section presents several existing IP FRR mechanisms. In Section 3, we introduce our proposed tunnel-based IP FRR mechanism with an illustrative example. The tunnel endpoint selection problem and our proposed algorithm for solving this problem are discussed in Section 4. In Section 5 and 6, we present evaluation methodology and results respectively. Finally, we conclude the paper in Section 7.

2. IP FRR STATE OF THE ART

In this section, we briefly introduce several IP FRR mechanisms are currently being investigated by the IETF.

2.1 Loop-free Alternate

In the Loop-free Alternate (LFA) scheme [4], when a *direct* neighbor of the repairing router has a native IGP path to the destination without traversing the protected network component, the repairing node can directly forward the affected traffic to that neighbor for achieving FRR when the failure of the protected network component is detected. A necessary condition for a neighbor to become a feasible candidate for achieving FRR is that this neighbor should not return the traffic back to the repairing router when the traffic is being delivered to the destination.

2.2 Tunneling

Apart from direct neighbors, a router that is more than one hop away from the repairing router can also be used for FRR. If none of the direct neighbors are feasible, the repairing router can send the traffic via an IP tunnel to a remote point in the network which has native IGP paths to the destination without traversing the protected link [5]. This remote router is called tunnel endpoint. In order to increase the failure protection coverage, the authors of [5] also proposed some advanced forwarding mechanisms such as directed forwarding in conjunction with IP tunnels.

2.3 Not-via Addresses

Not-via [6] uses special IP addresses assigned to each protected interface. The semantics of a not-via address is that "a packet addressed to a not-via address must be delivered to the router advertising that address, not via the protected component with which that address is associated". When a failure occurs, the repairing router encapsulates the packet to a not-via address of the protected interface. From the not-via address, the routers along the repair path can know to which next-hop they must deliver the packet in order to avoid traversing the failed interface. Not-via can always guarantee full failure protection, meaning that FRR is available for any destination under any single failure scenario, provided that there are not critical links whose failure will cause the network topology to be broken.

3. TUNNEL-BASED IP FAST REROUTE

A generic tunnel-based IP FRR mechanism is proposed in this paper. This mechanism shares some similarity with the one proposed in [5] since both use tunnel encapsulation for implementing the repair path. However, there are several key differences. First of all, our mechanism allows the use of dedicated tunnel endpoints for the repair paths to different destinations, while the existing mechanism uses only a single tunnel endpoint for all the affected destinations. Our proposed per-destination based scheme, which has also been used by LFA and also [7], provides higher flexibility in provisioning repair paths. Furthermore, in our mechanism, a tunnel endpoint always forwards the traffic *natively* to the final destination without relying on the additional direct forwarding mechanism [5], which cannot be naturally supported by conventional IP routers.

3.1 Operations and Illustrative Example

Our tunnel-based IP FRR mechanism allows repairing routers to be pre-configured with tunnel endpoints that are able to detour traffic from the protected link before reaching its final destination. The overall repair path consists of two shortest path segments: one from the repairing router to the tunnel endpoint (the tunnel) and the other from the tunnel endpoint to the final destination. Fig. 1 illustrates this rerouting operation.

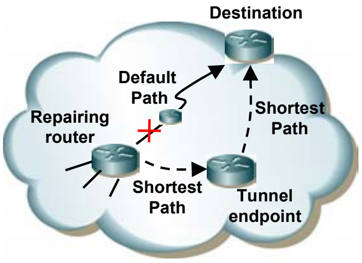


Figure 1. Repair path using the tunnel-based IP FRR mechanism

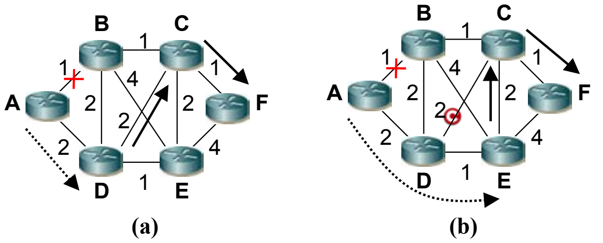


Figure 2. Example of the tunnel-based IP FRR mechanism

An example of our proposed scheme is illustrated in Fig. 2(a). Given the set of IGP link weights shown in the figure, the shortest path from router *A* to *F* is *A-B-C-F*. At router *A*, *D* can be selected as the tunnel endpoint for the repair path that protects link A-B with regard to the traffic towards *F*. In case link *A-B* fails, the repairing router *A* immediately re-routes the traffic away from *B* to the tunnel endpoint *D* via the IP tunnel (i.e. *A-D*). Next, *D* de-encapsulates the packets and forwards the traffic natively to the final destination *F* based on the conventional IP shortest path routing (i.e. the path *D-C-F*). However, if link *D-C* becomes congested due to the diversion of the affected traffic from the repairing router *A*, router *D* may not be a good choice of tunnel endpoint in the first place. To avoid potential post-failure congestion, router *E* may be used as the tunnel endpoint instead of *D*. In this case, as shown in Fig. 2(b), the traffic is effectively re-routed onto the repair path *A-D-E-C-F* without traversing link *D-C* that is prone to congestion. This example shows that our tunnel-based IP FRR scheme provides flexibility in optimizing post-failure network performance by judicious selection of tunnel endpoint. To achieve optimized post-failure traffic distribution with IP FRR, the network operator needs to obtain the following information *a priori* in order to perform optimized tunnel endpoint selection in an offline manner: the overall network topology including the IGP link weight setting, the forecasted traffic matrix and the distinct failure scenarios to be considered. This is similar to the input for the robust IGP traffic engineering [8].

4. TUNNEL ENDPOINT SELECTION

4.1 Problem Formulation

Given a network topology with configured IGP link weights, for each link to be protected, the repairing router may have multiple choices for selecting tunnel endpoint, and each could result in different post-failure network utilization. To minimize the possibility of creating post-failure network congestion, it is important to judiciously pre-determine the best tunnel endpoint such that the overall load distribution in the network after failure is balanced. We name this *IP FRR tunnel endpoint selection problem*. We focus on single link failures [1] but the proposed scheme can be easily adapted to router failures as well.

We now formally define the tunnel endpoint selection problem. Let $G=(V,E)$ represents a network topology with a set of routers V and a set of unidirectional edges E with $e(x,y)$ representing the link connected from router x to y. Based on the configured IGP link weights, the shortest path from router x to y is denoted by $x \rightarrow y$. Let $f_{x,y} \subseteq V \times V$ be the traffic that is sent from router x to destination y. Note that $f_{x,y}$ includes not only the traffic that is locally originated from x but also from the other routers in the network which must traverse x before reaching y. The tunnel endpoint selection problem is as follows:

For each adjacent link to be protected at each repairing router x, select a tunnel endpoint, denoted by $t_{x,y}$, for each affected destination y so that $f_{x,y}$ will be rerouted over $x \rightarrow t_{x,y} \rightarrow y$ when the protected link fails. An affected destination means that the shortest path from the repairing router to it involves the protected link. The ultimate goal is to avoid post-failure network congestion on the repair path due to careless selection of $t_{x,y}$.

We define Maximum Link Utilization (MLU) to be the utilization of the highest loaded link within the network. Under the failure of link $e(u,v) \in E$, let $\mu_{u,v}$ be the post-failure MLU after router u has rerouted the traffic for all the affected destinations via the selected tunnel endpoints. Since the tunnel endpoint selection is performed for each protected link independently due to single link failure protection, the optimization objective of the tunnel endpoint selection problem is to minimize the post-failure MLU for each of these scenarios, which is defined as

$$Minimize \quad \mu_{x,y} \qquad \forall e(x,y) \in E$$

4.2 Heuristic Algorithm

We propose a two-phase efficient algorithm for solving the tunnel endpoint selection problem.

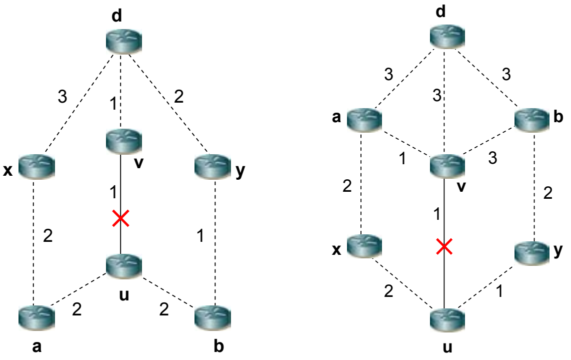
4.2.1 Phase 1: Feasible Tunnel Endpoint Filtering

Although any router in the network could be considered as tunnel endpoint candidate, some may cause forwarding loops and therefore are infeasible. The first step of our algorithm is to identify all the feasible tunnel endpoints for each protected link by its repairing router with regard to each affected destination. Let u and v be the head (i.e. repairing) and tail router of the link $e(u,v)$ to be protected respectively, d be the destination router, $w(u, v)$ be the IGP weight of link connecting from router u to v, and finally $dist(x, y)$ be the total IGP cost of $x \rightarrow y$. If a router is a feasible tunnel endpoint, two necessary conditions must be met:

***Constraint 1 (Not hidden behind repairing node)*:** For any router o in the network to be a feasible tunnel endpoint for u to reach destination d, u must *not* be on $o \rightarrow d$. That is:

$$dist(o, u) + dist(u, d) > dist(o, d)$$

***Example*:** In Fig. 3(a)[1], router *a* is an infeasible tunnel endpoint candidate for the protected link *u-v* with regard to the destination *d*. This is because once packets are de-capsulated at *a*, they will be attracted back to the repairing router *u* on their way to *d*. Router *b* is a feasible candidate since $b \rightarrow d$ does not involve the protected link.



(a) Constraint 1 **(b) Constraint 2**

Figure 3. Constraints for tunnel endpoint filtering

***Constraint 2 (Not hidden behind tail of protected link)*:** For router o in the network to be a feasible tunnel endpoint candidate for u to reach d, v must not be on $u \rightarrow o$, i.e.

$$w(u, v) + dist(v, o) > dist(u, o)$$

***Example*:** In Fig. 3(b), router *a* is an infeasible tunnel endpoint candidate for the protected link *u-v* with regard to the destination *d*. This is because the tunnel from the repairing router *u* to *a* still

[1] The dash lines in the figure indicate shortest IGP paths, which means the two routers are not necessarily connected physically.

traverses the protected link. Router *b* is a feasible candidate since *u* → *b* does not involve the protected link.

4.2.2 Phase 2: Tunnel Endpoint Selection

Given the set of feasible tunnel endpoints identified in the first phase, the second phase of the algorithm is to select the best tunnel endpoint such that the overall post-failure MLU under the considered link protection scenario is minimized. The basic idea of the second phase is to first identify all the affected destinations for each of the adjacent links to be protected. Then, for each of these destinations, select the best feasible tunnel endpoint in a greedy fashion with the objective to minimize the corresponding MLU assuming the failure of the protected link. The detailed steps of the algorithm are as follows.

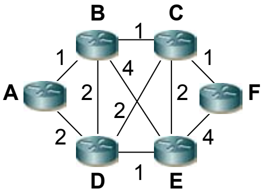
```
Input 1: A set of feasible tunnel endpoints to each affected
destination for each protected link.
Input 2: Network topology and traffic matrix.
Step 1:  Set $\Omega$ to be the current network (normal) status.
Step 2:  For router $x$, consider a directly attached link to be
         protected.
Step 3:  Identify all destinations $y \in V$ where the shortest paths $x$
         $\rightarrow y$ traverse the protected link. Then, remove their
         traffic $f_{x,y}$ from $x \rightarrow y$.
Step 4:  Sort all the destinations in descending order according
         to their associated traffic volume $f_{x,y}$.
Step 5:  For each destination $y$ in that order,
         if there exist feasible tunnel endpoints for $y$, then
           – try to route $f_{x,y}$ to the destination via each of the
             feasible tunnel endpoints independently and records
             the corresponding post-failure MLU.
           – select the one that results in the least MLU as the
             tunnel endpoint $t_{x,y}$.
           – update the network by routing $f_{x,y}$ over $x \rightarrow t_{x,y} \rightarrow y$.
Step 6:  Restore the current network status to $\Omega$.
Step 7:  Go to step 3 to consider the next adjacent link to be
         protected until all the adjacent links of router $x$ have
         been processed.
```

4.3 Illustrative Example of the Algorithm

We illustrate the operations of our algorithm using an example in Fig. 4. We consider router *A* in the network to be the repairing router of its two directly attached links *A-B* and *A-D*. *A*'s routing table in the normal state is also shown.

The algorithm starts with identifying all the feasible tunnel endpoints to each affected destination according to the two filtering criteria. Fig. 5 shows all feasible tunnel endpoints for each destination at router *A* to protect each of its adjacent links. This procedure repeats at each router in the network for every destination.

The next step of the algorithm is tunnel endpoint selection to achieve post-failure load balancing. Given the set of feasible tunnel endpoints, the algorithm proceeds as follows. First of all, consider an adjacent link of the repairing router to be protected, e.g. link *A-B* by router *A*. In this case, traffic for *B*, *C* and *F* is affected as their shortest paths traverse the link. Given $f_{A,B}$, $f_{A,C}$ and $f_{A,F}$, the algorithm removes these traffic from the network and then performs a sorting according to their traffic volume. Assuming that the sorting order is $f_{A,F}$, $f_{A,C}$, $f_{A,B}$. For the first destination in that order (i.e. *F*), the algorithm selects between *D* and *E* as the tunnel endpoint. By trying each of these tunnel endpoints one at a time over the corresponding repair paths (i.e. *A* → *D* → *C* → *F* and *A* → *D* → *E* → *C* → *F* respectively), the one that results in the least MLU is selected. If, for example, *D* is selected, $f_{A,F}$ will be rerouted in the network over *A* → *D* → *C* → *F*. Given this updated network topology, the next destination *C* is tried using the above procedure until the last destination *B* has been considered. Fig. 5 shows an example of the final tunnel endpoint selection result.



Dest	Next Hop
B	*B*
C	*B*
D	*D*
E	*D*
F	*B*

Figure 4. Network topology and routing table of router *A*

Protected link	Dest	Feasible tunnel endpoints	Selected tunnel endpoint
A-B	*B*	*D or E*	*E*
	C	*D or E*	*E*
	F	*D or E*	*D*
A-D	*D*	*B or C or F*	*C*
	E	*B or C or F*	*B*

Figure 5. Example of tunnel endpoint selection

5. EVALUATION METHODOLOGY

5.1 Network Topology and Traffic Matrices

Our evaluation uses the topologies and traffic matrices from two real operational networks: GEANT [9] and Abilene [10]. The GEANT network topology consists of 23 PoP routers and 74 unidirectional links, while Abilene consists of 11 routers and 28 unidirectional links. We use real traffic matrices for these two operational networks which are available from [11].

5.2 Performance Metrics

The following performance metrics are considered:

5.2.1 Fast Reroute Coverage

It indicates the coverage of failure protection by the proposed tunnel-based IP FRR mechanism. We use the following metrics defined in [3]:

- **(FC-1) The percentage of links which can be fully protected for all destinations**: a link is said to be fully protected for all destinations only if every affected destination has at least one feasible tunnel endpoint.
- **(FC-2) The percentage of destinations which can be fully protected for all single link failure scenarios:** a destination is said to be protected for all link failures if there exist at least one feasible tunnel endpoint for every link failure that affects the destination.

- **(FC-3) For all destinations and for all link failures, the percentage of the total potential failure cases which are protected.** This examines the overall "best effort" protection.

5.2.2 Post-Failure Maximum Link Utilization

For simplicity we assume each network link has equal chance to fail, but with no simultaneous failures of multiple links. We then consider the worst-case (i.e. highest) post-failure MLU among all the single link failure scenarios. Note that if a link cannot be fully protected for all destinations, we assume IP re-convergence will handle this type of link failure instead. Therefore, the worst-case post-failure MLU could be the result of our tunnel-based IP FRR mechanism or IGP re-convergence.

5.3 Approaches for Performance Comparison

We consider the following approaches in our evaluation of post-failure network utilization:

- **IGP-RCVG:** This approach relies on IGP re-convergence to recover routing failures. It is thus a basic and reactive approach that handles link failures without considering IP FRR.
- **FRR-G:** This approach adopts the tunneling mechanism for IP FRR in conjunction with our proposed tunnel endpoint selection algorithm. It is thus the approach that considers both fast routing failure recovery and post-failure load balancing.
- **FRR-R:** This approach is similar to the FRR-G except that the tunnel endpoint selection is purely done *randomly*. It can be regarded as the approach that aims to achieve fast recovery from routing failure only, without considering post-failure load balancing. Note that, as random selection may produce results with different worst-case MLU, we take an average of 10 independent trials.

5.4 IGP Link Weights

For each of the approaches above, the following IGP link weights are used in our evaluation:

- **Actual:** Actual link weight in real operational networks.
- **Uniform:** Homogeneous link weight.
- **InvCap:** Link weights proportional to inverse capacity.
- **TE-Optimized:** Taking as input a traffic matrix, link weight is optimized with the objective of minimizing the network cost, as proposed by Fortz & Thorup [8]. We use the TOTEM toolbox [11], which has the implementation of their local search heuristic, to generate the link weight.

6. RESULTS

6.1 Fast Reroute Coverage

For the TE-optimized link weight scenario, we further tested three traffic matrices with different levels of traffic intensity under normal network conditions (i.e. low, medium and high). Therefore, the link weights used for the three traffic scenarios may be different.

From the results in Table 1 for the GEANT network, we see that IGP link weight configuration plays an important role in influencing the degree of FRR coverage. This suggests that *the FRR coverage could be improved by optimizing IGP link weights*. As far as *FC-1* is concerned, our tunnel-based IP FRR mechanism can protect all affected destinations from most of the link failures. The percentage of destinations that can be protected by all link failures (*FC-2*) as well as *FC-3* is also high in general. These results show that the proposed tunnel-based IP FRR mechanism is effective.

Table 1. Fast Reroute Coverage for GEANT

Link weight	FC-1	FC-2	FC-3
Actual	94.6%	69.57%	98.61%
Uniform	100%	100%	100%
InvCap	97.3%	60.87%	98.22%
TE-Optimized (low)	98.65%	95.65%	99.8%
TE-Optimized (med)	98.65%	39.13%	97.23%
TE-Optimized (high)	100%	100%	100%

Table 2. Fast Reroute Coverage for Abilene

Link weight	FC-1	FC-2	FC-3
Actual	78.57%	27.27%	84.55%
Uniform	100%	100%	100%
InvCap	100%	100%	100%
TE-Optimized (low)	92.86%	63.64%	96.36%
TE-Optimized (med)	92.86%	72.73%	97.27%
TE-Optimized (high)	89.29%	54.55%	94.55%

For the TE-optimized link weight scenario, although the same algorithm is used, the results based on different traffic matrices can lead to different degree of FRR coverage. This implies that *there may be a tradeoff between traffic performance and effectiveness of IP FRR in which achieving near-optimal network performance comes at the expense of low degree of failure protection by IP FRR*.

For the Abilene network, we found that the results are similar to those of the GEANT network except that the degree of FRR coverage is in general lower, mainly due to small average node degree of the topology which reduces the number of feasible tunnel endpoints.

We also observed an interesting result from both tables that homogeneous link weight can always achieve full FRR coverage in general. This suggests that *there exist at least a set of link weights that can always achieve full FRR coverage by using the proposed tunneling mechanism*. An analytical proof for this is given in the Appendix.

For the results in Table 1 and 2, we see that full FRR coverage cannot be achieved under some network configuration scenarios. We believe that there are two primary reasons that can influence the FRR coverage: IGP link weights and topology connectivity. For IGP link weights, we have already demonstrated its effects in Table 1 and 2. On the topology connectivity side, a straightforward approach is to add more links in order to enrich the overall network connectivity. A natural hypothesis is that adding links to the network increases the chance in finding feasible tunnel endpoints, thereby improving the FRR coverage.

By using BRITE [2], we generate a large-scale network topology with 50 routers and 200 unidirectional links. For the purpose of demonstration, we assume that link capacity is randomly generated and InvCap link weight setting is used. Links are added to the base network following the Waxman's model, i.e. links are added between routers that are closest to each other if there were not any link exist.

Table 3. Network topology expansion

Network Topology	FC-1	FC-2	FC-3	Worst-case MLU
Base	99.5%	98%	99.96%	172%
+ 6 links	99.03% ⬇	86% ⬇	99.71% ⬇	134% ⬆
+ 12 links	99.06% ⬇	80% ⬇	99.59% ⬇	93% ⬆
+ 16 links	99.07% ⬇	84% ⬇	99.67% ⬇	93% ▬

Table 3 shows the FC-1, FC-2, FC-3 and worst-case MLU with different number of links added onto the base network topology. An interesting result is that when adding 6, 12 or 16 links to the network, all the values of FC-1, FC-2 and FC-3 decrease compared to those for the base topology. This reveals that *adding more links to the network does not guarantee the improvement of FRR failure protection coverage. Instead, this could make the performance even worse*. These results refute the original hypothesis. Nevertheless, the worst-case MLU is improved as more capacity is available in the network.

6.2 Post-Failure Load Balancing

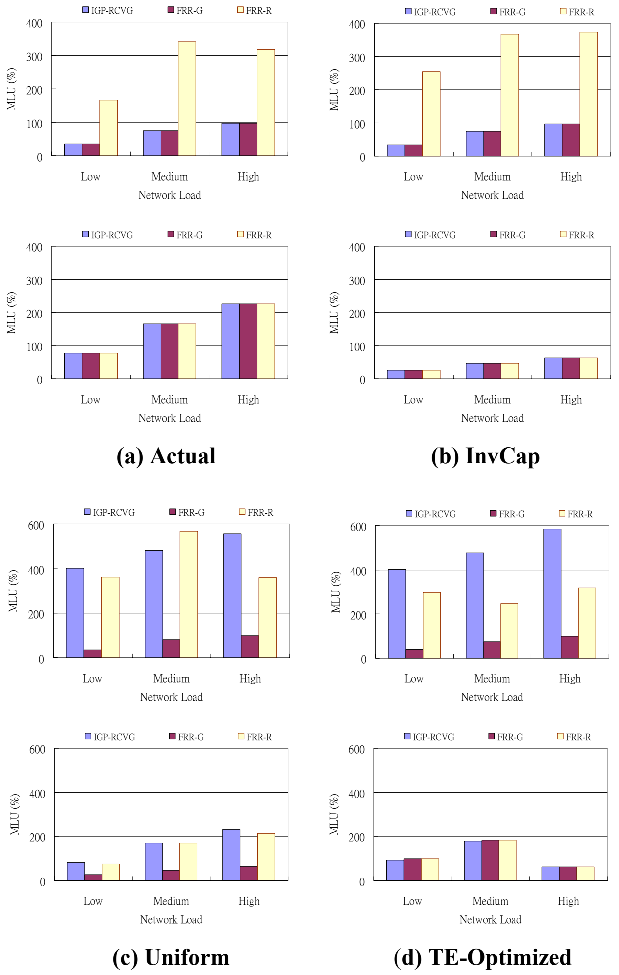
Having evaluated the FRR coverage achieved by the proposed tunnel-based IP FRR mechanism, we proceed to investigate how different approaches in Section 5.3 perform in terms of post-failure network utilization. Fig. 6 shows the worst-case MLU among all the link failures.

An overall picture of Fig. 6 indicates that the FRR-G approach performs much better than the IGP-RCVG approach. This means that using IP FRR via tunneling together with judicious tunnel endpoint selection can achieve significant performance improvement over the traditional IGP re-convergence. The gain includes not only very high FRR coverage for failure protection but also minimized possibility of experiencing post-failure network congestion. On the other hand, imprudent tunnel endpoint selection can cause severe congestion after the affected traffic is diverted onto the randomly selected tunnel endpoint. As a result, packets may still be discarded even though routing failures can be recovered rapidly, thereby not able to guarantee comprehensive QoS assurance.

However, under some scenarios in Fig. 6, the FRR-G approach has the same performance to the IGP-RCVG or even the FRR-R approach. This is partly due to the reason that, under failure of some links, there does not exist any feasible tunnel endpoint for some destinations. As a result, FRR is not used when *any* of these link failures occurs and the traditional IGP re-convergence takes place instead. However, we observed that one of these link failures has caused the highest utilization. This explains why the worst-case MLU of both approaches is the same as they both account for the same highest utilization based on the traditional IP re-convergence.

Another interesting result is shown in Fig. 6(d) Abilene network. For the TE-optimized link weight, the worst-case MLU based on the high-loaded traffic matrix is surprisingly lower than those based on the low- and medium-loaded ones. In general, the higher the traffic load is, the highest the link utilization. However, this abnormal phenomenon implies that the link weight that is well optimized only for the normal network condition could perform poorly under link failures with IP FRR.

Based on the evaluation results in Table 1, 2 and Fig. 6, we conclude with the following observations. (i) The proposed tunnel-based IP FRR mechanism with judicious tunnel endpoint selection can achieve both fast routing failure recovery and post-failure load balancing; (ii) Imprudent tunnel endpoint selection can easily cause severe congestion after failures and therefore reduces the effectiveness of IP FRR. Therefore, it is not advisable to deploy IP FRR only without considering how to control it to achieve post-failure load balancing; (iii) Link weights that are optimized for conventional traffic engineering under failure-free conditions may lead to lower FRR coverage and poor post-failure network utilization after affected traffic is re-routed onto the repair paths.



(a) Actual **(b) InvCap**

(c) Uniform **(d) TE-Optimized**

Figure 6. Post-failure network utilization under various IGP link weights (top: GEANT, bottom: Abilene)

7. CONCLUSION

Most of the current IP fast reroute mechanisms only focus on achieving rapid recovery from routing failures. However, they do not consider the network performance after failure when deployed in which case post-failure network congestion may be incurred. This paper investigated a tunnel-based IP fast reroute mechanism which takes this issue into account. An efficient tunnel endpoint selection algorithm is proposed to achieve post-failure load balancing. Our evaluation results based on real operational networks reveal that the tunneling mechanism with judicious selection of tunnel endpoint can achieve high fast reroute coverage and improve post-failure load balancing. We also found that IGP link weight plays an important role in influencing the overall failure coverage, which facilitates network operators to intelligently configure the IP routing logic (e.g. setting appropriate IGP link weights) to achieve maximum fast reroute coverage.

APPRENDIX — Proof for full failure coverage with homogenous link weight setting

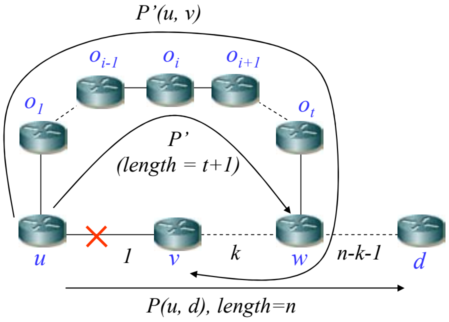


Figure 7. Full failure coverage with hop count routing

For simplicity we consider the scenario that all IGP link weights are set to 1 (hop count routing). As shown in Fig. 7, router u and router v are the head router and tail router of the protected link l respectively, and l is on the shortest IGP path from u to a specific destination d. The shortest IGP path from u to d in the normal state is denoted by $P(u, d)$ whose length is n hops. First we find an alternative shortest IGP path $P'(u, v)$ from u to v that does *not* involve the protected link l. If no such a path exists, then l is a critical link and naturally there is no feasible solution to protecting l. We consider the general case that $P'(u, v)$ has overlapping path segment with $P(u, d)$. Assume the first merging point of the two paths (from the viewpoint of $P'(u, v)$) is node w and the length of the overlapping segment is k. If $k=0$ it means $P'(u, v)$ and $P(u, d)$ are link-disjoint. Now we consider the path segment between u and w on $P'(u, v)$ which is denoted by P'. We denote the intermediate router sequence as $o_1, o_2, \ldots o_i, \ldots o_t$ ($t>0$) on P' and in this case $dist(P')=t+1$[2].

If t is an odd number, then router $o_{(t+1)/2}$ is a feasible tunnel endpoint candidate for u. This is because: $dist(o_{(t+1)/2}, u)+dist(u,d)=(t+1)/2+n=>dist(o_{(t+1)/2}, d)=(t+1)/2+n-k-1$, given $k\geq 0$. This means that constraint 1 (not hidden behind repairing node) is satisfied. Moreover, $w(u,v)+dist(v,o_{(t+1)/2})=1+k+(t+1)/2>dist(u,o_{(t+1)/2})=(t+1)/2$, given $k\geq 0$. This means constraint 2 (not hiding behind tail of protected link) is satisfied.

The case that t is an even number is a little more complicated. If we consider router $o_{t/2+1}$ we have $dist(u, o_{t/2+1}) = dist(o_{t/2+1}, u) = t/2+1$ and $dist(o_{t/2+1}, v) = t/2$. First, $dist\ (o_{t/2+1}, u) + dist\ (u, d) = t/2+1+n > dist\ (o_{t/2+1}, d) = t/2+n-k-1$, given $k\geq 0$. This means constraint 1 is satisfied. Next, $w(u, v)+dist(v, o_{t/2+1})=1+k+t/2\geq dist(u, o_{t/2+1})=t/2+1$. This means constraint 2 is satisfied only if $k>0$. In case $k=0$ then the two distinct paths $\{u, o_1, \ldots, o_{t/2+1}\}$ and $\{u, v(w), o_t, \ldots o_{t/2+1}\}$ have equal length as far as constraint 2 is concerned. As a result, in order to make router $o_{t/2+1}$ a feasible tunnel endpoint candidate for u, u should use the shortest path $\{u, o_1, \ldots o_{t/2+1}\}$ to tunnel to $o_{t/2+1}$ instead of passing through the protected link (u, v).

8. ACKNOWLEDGMENT

This paper is partially supported by the European Union IST FP6 AGAVE project.

9. REFERENCES

[1] G. Iannaccone et al., "Analysis of Link Failure in a Large IP Backbone," Proc. ACM IMW, 2002.

[2] BRITE topology generator: http://www.cs.bu.edu/brite/

[3] M. Shand and S. Bryant, "IP Fast Reroute Framework," draft-ietf-rtgwg-ipfrr-framework-08, February 2008.

[4] A. Atlas et al., "Basic Specification for IP Fast Re-route: Loop-free-Alternates," draft-ietf-rtgwg-ipfrr-spec-base-11, February 2008.

[5] S. Bryant et al., "IP Fast Reroute Using Tunnels," draft-bryant-ipfrr-tunnels-03, November 2007.

[6] M. Shand et al., "IP Fast Reroute Using Not-Via Addresses," draft-ietf-rtgwg-ipfrr-notvia-addresses-02, February 2008.

[7] S, Nelakuditi, "Fast Local Rerouting for Handling Transient Link Failures," IEEE/ACM Transactions on Networking, April 2007.

[8] B. Fortz and M. Thorup, "Robust Optimization of OSPF/IS-IS Weights," Proc. INOC, 2003.

[9] The GEANT topology: http://www.geant.net/upload/pdf/GEANT_Topology_12-2004.pdf

[10] The Abilene topology and traffic matrices dataset: http://www.cs.utexas.edu/~yzhang/research/AbileneTM/

[11] The TOTEM project: http://totem.info.ucl.ac.be/

[2] Although in the figure all the indicated paths are uni-directional, due to the fact that all IGP link weights are set to 1, the paths in both directions are symmetric.