Fast Scoped Rerouting for BGP

Roland Bless, Götz Lichtwald, Markus Schmidt
and Martina Zitterbart
Institute of Telematics
University of Karlsruhe, Germany
Email: {bless, lichtwald, mschmidt, zit}@tm.uka.de

***Abstract*— This paper[1] presents an approach to improve inter-domain connectivity in the Internet. This novel concept deploys inter-domain routing functions on two different time scales. The innovative Fast Scoped Rerouting approach operates on a fine granular time scale while regular BGP is used on a coarse granular time scale. The overall concept intends to provide fast recovery from failures and to reduce the amount of globally visible BGP update messages. It also provides an alternative path in case of failure. Thus, this novel approach improves the Internet's ability to derive a coherent view of its topology.**

I. Introduction

The currently deployed inter-domain routing protocol in the Internet is the Border Gateway Protocol (BGP) [1], [2]. It provides inter-domain connectivity and considers complex rules (policies) to influence the paths that are propagated to peers. At the time BGP was developed, policy-aware connectivity was the main objective. Today, however, the Internet user community requires not only a stable connectivity but also fast recovery from failures. Those requirements can not be fulfilled by BGP [3], [4] entirely. The scalability of the current Internet depends on several parameters, like the number of Autonomous Systems (ASes) or distinct routing table entries. Those parameters affect scalability as they determine the amount of exchanged information and the required resources, i.e., CPU time and memory usage, to fulfill the routing task. Every new AS adds at least one entry to the routing table. But because multi-homing is increasingly used, it is usually much more than only one entry. Another, not negligible, criterion for scalability is the dynamic of the network [3] – especially the number of BGP updates, i.e., withdrawals and announcements of routes [5].

The frequency of BGP updates is one of the most serious problems of the current Border Gateway Protocol. In [5]–[13] many different reasons for the occurrence of BGP updates are given. Among of them are: Router configuration errors (so called human errors), transient short-time physical and data link problems, software bugs, problems with leased lines (electrical timing issues that cause false alarms of disconnect) or short-time router failures.

All those reasons have one thing in common: The generated BGP update is, strictly speaking, unnecessary, because it could have been avoided due to the temporary and short-time nature of the failure. Nevertheless, each time such a failure occurs, a BGP update has to be issued. Due to the fact that BGP propagates every update message globally, the whole Internet is stressed even by a single mis-configuration. According to [6] the main reason for routing instability are mis-configurations.

Looking at an enormous mass of BGP updates populating the Internet [10], [11] a mechanism is needed to reduce the total amount of BGP updates. Not only the load of router CPUs and the network is affected by those updates, but also the Internet has almost no chance to reach a consistent view at a single point in time.

Several recently proposed approaches try to alleviate BGP update storms. Most of them fix only a single BGP problem and extend BGP in a patchwork manner. The most related approaches to our novel concept are described in section II.

The approach presented in this paper limits updates—in first instance—to those BGP peers that are directly affected by the current network change (e.g., link failure). Furthermore this novel approach provides an alternative path to substitute a broken AS path. During the first reaction to the failure only peers that are inevitably affected by the failure are stressed with update messages. This approach is called *Fast Scoped Rerouting (FaSRo)*, because it routes around the failure involving only a few peers to establish an alternative AS path. This reduces the total number of routing messages and accelerates convergence time [14], [15]. In contrast to that, BGP propagates a link failure in inter-domain connectivity globally.

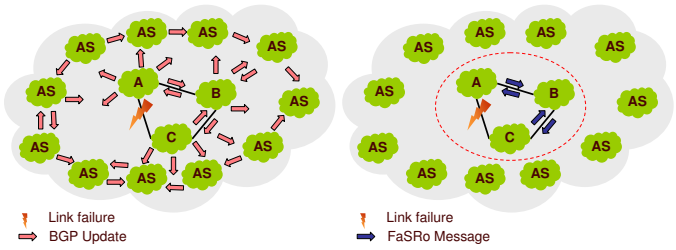
This paper is structured as follows: Section II contains a closer look at approaches that try to alleviate the problems of too many BGP update messages. Section III presents our novel FaSRo approach to improve the stability of inter-domain connectivity. Section IV provides first simulation results. Finally, section V gives a conclusion and an outlook on future work.

II. Problems of recent alleviation approaches

Looking at the current BGP protocol for almost every change in the network a globally visible BGP update message is issued (cf., Fig. 1(a)). Contrary to that the novel FaSRo approach described in this paper provides a limited propagation scope for a temporary failure in the inter-domain connectivity (cf., Fig. 1(b)).

Basically three concepts can be distinguished that try to alleviate the problem of too many BGP update messages:

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



(a) BGP update ripples (b) FaSRo propagation

Fig. 1. Comparison of update propagation between BGP and FaSRo

Graceful Restart [16], Route Flap Damping [17] and the recently proposed NOPEER attribute [18].

A. Graceful Restart

The *Graceful Restart Capability* [16] introduces a new transitive BGP attribute that describes the capability of the BGP router to convey routing convergence information to its peers. This information is actually propagated via the *END-of-RIB* marker.

The main idea of this concept is that a Graceful Restart capable BGP router is able to preserve its forwarding information during its restart process.

This mechanism prevents route flapping due to holding back the re-computation of the routing table during the restart process. The computation of the routing table is issued as soon as the *END-of-RIB* marker is received. In case the peer of the restarting router is not Graceful Restart capable the re-computation starts immediately.

B. Route Flap Damping

A further approach to alleviate the inter-domain instability problem is *Route Flap Damping* [17].

This mechanism keeps a penalty value for each peer and per destination. This penalty value is increased every time a route change announcement, i.e. a BGP update message, is received. The penalty value decays exponentially. During the penalty phase, updates are ignored that would lead to network improvements (e.g., a broken path is recovered). This is considered as a big disadvantage of Route Flap Damping [19], [20].

In [19] an example is given where a route flaps within a two minute interval. This behavior causes a Cisco router to suppress this route on the third flap for more than 28 minutes, if the Cisco router uses the recommended set up values [17] for penalizing a route flap.

C. NOPEER-Attribute

This approach [18] suggests a *scope control BGP community* to allow an origin AS to determine to which extent a route is propagated externally. The boundary of the propagation scope has to be determined a priori. Thus, it is not possible to react on sudden disrupting network changes.

This concept addresses network issues like limited transit services where advertisements are restricted to certain transit providers and various forms of selective transit in a multi-homed environment.

D. Rating

None of the listed approaches is able to react dynamically with respect to the scope that is used for the propagation of updates. Graceful Restart notifies peers about converged routing information via the END-of-RIB marker but provides no mechanism to alleviate BGP update storms. The NOPEER attribute needs a predefined scope to limit the route propagation and has no dynamic mechanism to adapt the propagation scope in case of a failure. Route Flap Damping is different from Graceful Restart and the NOPEER attribute concept, because it reacts dynamically on route flaps. But Route Flap Damping is considered to be far too strict concerning network protection [19], [20], because even good news are blocked during the damping phase.

III. The Fast Scoped Rerouting Approach

A. Basic Concept

The main goal of this novel approach is to improve the inter-domain routing stability and to reduce the convergence time. The basic idea is to limit the notification scope of updates and to switch to an alternative AS path. FaSRo covers only one part of the inter-domain routing process. It achieves a fast reaction to a short-time problem for the trade-off of temporarily installing a non-optimal route from a global point of view. Thus, global propagation of a route change is the second part (usually handled by normal BGP updates) that restores globally optimal routes again. This is especially important if the problem is not temporary but rather persistent. Therefore, the overall concept is based on two time scales to propagate changes:

- The *fine granular time scale* process is used to handle AS path changes, e.g., a broken link. A local scoped reaction takes place that establishes an alternative path (see section III-B) to substitute the broken one.
- On the *coarse granular time scale* BGP takes control of the broken link. This is the case if the failure stays persistent for a certain period of time. Then BGP updates are issued and new routes are calculated by BGP.

For the fine granular time scale process it is assumed that peers along the alternative path are FaSRo capable. If this is not the case, routers can only speak BGP. Consequently, the advantages of the new concept cannot be utilized in this case, but connectivity will not be broken by this fact. No other side-effects will be observed with respect to connectivity compared to regular BGP operation.

In the following, the term “link failure” is used to denote the total loss of inter-domain connectivity between two adjacent ASes (which may however be provided by a set of several redundant physical links).

To illustrate the overall concept of this novel approach an example is provided in Fig. 2. The upper part depicts the link failure scenario between AS A and D. The lower part depicts

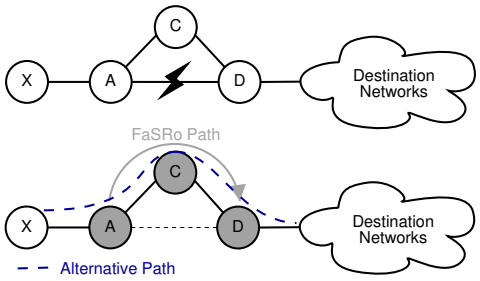


Fig. 2. Example

both the alternative path and the so-called *FaSRo path* that is established to substitute the broken link between A and D.

A link failure is—in the first instance—handled by FaSRo. Fig. 3 illustrates the way a failure is processed by using the state machine of the FaSRo process. Any message that would normally reach the Finite State Machine (FSM) of BGP is now redirected to state *S1* of the FaSRo process, which is an extension of the BGP FSM. State *S1* decides whether the message has to be processed by BGP, which is the case for so called `KEEPALIVE` messages and normal BGP updates or if the incoming message is forwarded to the FaSRo process, which is the case for link failures. At this point the FaSRo process starts handling the failure. The following three issues are described in detail:

- A failure, i.e., a broken link, is detected (section III-A.1)
- The broken link is recovered (section III-A.2)
- The failure stays persistent (section III-A.3)

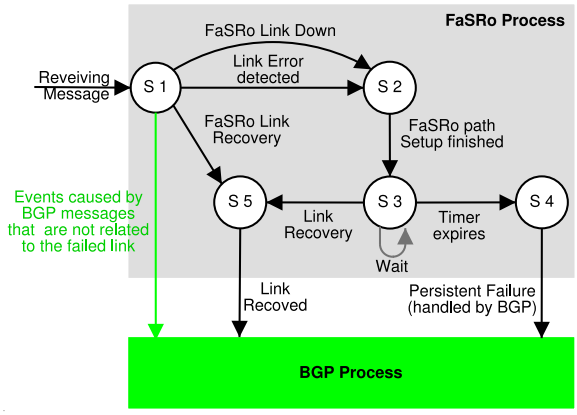


Fig. 3. FaSRo State Machine

1) Link Failure: Fig. 3 illustrates the behavior of the FaSRo process in case of a failure. As soon as the FaSRo process recognizes the failure of a link (this information is forwarded from state *S1* to state *S2*) two actions take place:

- The FaSRo process disables all those paths in its Forwarding Information Base (FIB) that use the broken link. Section III-B contains a detailed description how paths are disabled and how FaSRo paths are determined.
- The FaSRo process issues a message to its peers. These peers—from the failure detecting routers point of view—are members of the FaSRo path that is established to substitute the failed link. After this action is completed, state *S3* of Fig. 3 is reached.

As long as the FaSRo process has control of the error processing, BGP is not aware of the link failure. It is necessary to keep the BGP FSM unaware in order to avoid BGP issuing update messages.

The propagation of the link failure (starting at A cf., Fig. 2) is an AS hop-by-hop process along the FaSRo path (cf., Fig. 2 the hops are: A, C and D) that is set up by the FaSRo process. This AS hop-by-hop process terminates at the other end (cf., Fig. 2 AS D) of the FaSRo path that substitutes the link failure.

In the FaSRo context, compensation of a failure consists of two parts:

- Notification and establishment of the FaSRo path.
- A period T throughout which the corresponding state information is kept.

For T a suitable value has to be determined. In [6] it was observed that most of the temporary errors are recovered within 10 minutes. Thus, this period is suggested as a reasonable value of T.

2) Path Recovery: If the formerly broken link recovers within period T state *S5* (cf., Fig. 3) is reached. On entering state *S5* two actions are performed:

- The FaSRo process restores those paths that were marked as disabled due to the temporary failure of the link.
- The next peer that is part of the FaSRo process is notified about the recovery.

Every peer performs both actions. This procedure is executed along the FaSRo path that was established in order to substitute the temporary link failure. The whole process is finished after all routers on the FaSRo path have been notified and the control is given back to BGP (*BGP process* of Fig. 3).

3) Persistent Error: In case the duration of the link outage exceeds the so called *FaSRo Timer* T (as depicted in Fig. 3, state *S3*) it is assumed that the problem is not of temporary nature (state *S4*). Consequently, the behavior is switched back to normal BGP operation (*BGP process*). Thus, the link failure has to be propagated via BGP. At this point in time BGP update messages will be issued and a new route calculation will be performed.

B. Alternative Path

In the following, it is described how the FaSRo path is determined and which peers are members of this path.

The *Local Routing Information Base* (Loc-RIB [1], [2], [21]) contains all currently selected best paths. The FaSRo process on the BGP router that detects the failure selects an alternative path, based on its Adj-RIBs-In (RIBs composed from incoming routing information of adjacent peers). It notifies its next peer on this path about the fact that it is now member of the FaSRo process. For propagation among the peers two possibilities exist: Extending the Border Gateway Protocol, which is the preferred approach, or, developing a new protocol.

The FaSRo instance on every BGP peer along the FaSRo path will execute the following actions:

(i) For every destination network prefix d that is affected by the link failure between ASes A and D, determine an alternative path.
(ii) In order to provide a substitution for the broken link a FaSRo path has to be established.

Having a closer look at step (i) of the FaSRo process the following operations have to be performed:

- Search the Loc-RIB $\mathcal{L}$ for the next network prefix d whose route $p_{Best}(d)$ is affected by the link failure between A and D, i.e., $(A, D) \in p_{Best}(d)$. Mark this route as *invalid*.
- Determine the alternative path $p_{Alt}(d)$ to destination network d from Adj-RIBs-In $\mathcal{A}$. The alternative path must fulfill the condition that a sub-path of the alternative path $p_{Alt}(d)$ contains A and D as ASes.
  From the set of possible alternative paths to all affected destinations the sub-path with the shortest substitution between A and D is selected. This substitution is termed *FaSRo Path* p_{FaSRo}. Every further affected destination network prefix d is re-routed via p_{FaSRo}.

The task of step (ii) of the FaSRo process is to establish the FaSRo path. The path p_{FaSRo} is traversed hop by hop:

- Send the *FaSRo Link Down* message to the next hop of the FaSRo path (the next hop is determined by the AS path). This message notifies the peer about the fact that it now participates in the FaSRo process and advises it to establish the FaSRo path directed to D.
- For all destinations d whose paths included the link of the adjacent ASes A and D: forward all packets along the FaSRo path p_{FaSRo}.

Establishing only a single FaSRo path for all affected destinations is reasonable because only short-time outages are handled by FaSRo. Setting up an individual FaSRo path for every destination network prefix d would cause too much effort with respect to the routers' CPU and network load. Thus, all affected traffic is forced to take the FaSRo path which may cause policy violations in some cases. This, however, is acceptable since it is limited to the FaSRo timeout period.

Every provider using the FaSRo process profits in case of failure by other providers temporarily taking over its traffic and vice versa. Those providers that do not want to take over other ISP's traffic simply do not apply the FaSRo process.

C. Loops

In order to guarantee the property of loop-free paths after convergence two main cases have to be considered:

1) Can loops be created outside the FaSRo scope if the FaSRo process is started?
2) Can loops occur inside the FaSRo scope?

Concerning the first point the Border Gateway Protocol is still responsible to avoid loops. Loops can not occur because the AS_PATH field—containing the traversed Autonomous Systems—is transmitted on every route change announcement.

For the second point, two further cases have to be distinguished:

1) *Scenario:*
   A link failure was detected, but this information has not yet been propagated to all the peers that are part of the FaSRo path (so the set up process is still in progress).
   Conclusion:
   In this case packets are still sent via the old route and may cycle until the FaSRo path is completely established. To illustrate this behavior assume that AS C form Fig. 2 has as default path <A, D> to deliver packets to a destination network. After the link failure the FaSRo path is established among the ASes A, C and D. As long as C is not notified that it participates the FaSRo process C delivers packets to A and A forwards them again to C. This might cause packet loss if the packets' Time To Live value is counted down to zero and are thus discarded. This may only happen during the convergence time of the routing protocol, what—in fact—is nothing unusual, because all routing protocols can create transient loops during their convergence time.
2) *Scenario:*
   The information concerning the link failure has been propagated to all peers on the FaSRo path.
   Conclusion:
   If a loop occurs the loop must have been present before, i.e., BGP routes had not converged. Because every router on the FaSRo path is aware of its function as temporary re-routing peer. The FaSRo process can not cause the set up of a loop. No loops can occur due to the combination of BGP and FaSRo (for proof see section VI).

D. Improvements

Using the FaSRo process can improve the overall inter-domain routing situation. Due to the application of two time scales only little network and router CPU load is generated in the first instance, i.e., when no BGP mechanism is used.

As described above, failures and mis-configurations affect at first only a limited number of BGP routers (cf., Fig. 1(b)). The same applies to the restoration of a link respectively to the correction of a mis-configuration.

So failures—of whatever kind—affect only a reduced scope compared to the normal BGP failure propagation mechanism. This reduction allows fast recovery and makes it possible to accelerate the convergence time, because there is no need to re-negotiate policy-conforming routes that would replace the broken link. Using BGP would require a re-calculation of the routing table and negotiation of new paths so that those are conforming to the current policies.

Thus, in the first instance only a fast switch-over from the broken path is provided. It is switched back to the former path when the broken path is recovered. This behavior provides more inter-domain routing stability, because the amount of globally visible BGP updates is drastically reduced. In contrast to route flap damping FaSRo does not need to block BGP updates from a particular peer concerning a certain destination

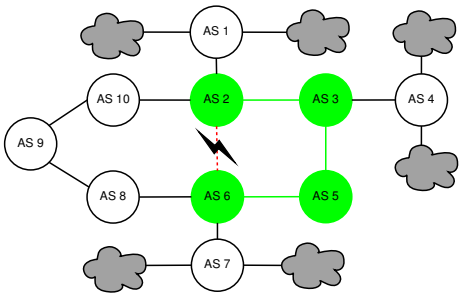


Fig. 4. Topology used for evaluation

to protect the routers' CPU and to reduce the network load. FaSRo is a CPU and network load friendly mechanism and does not consume a lot of resources. The inter-domain routing topology becomes more stable and consequently the Internet has the opportunity to converge to a coherent view.

Note that FaSRo cannot help if a complete AS fails. As this case is rather unlikely, FaSRo was not designed for such an event.

IV. First Evaluation

In order to prove that the FaSRo concept works as expected a simulation was set up using the event discrete simulation kit *OMNeT++* [22]. The whole FaSRo concept was simulated on AS level basis.

For the simulation a lightweight BGP engine was implemented. This engine only issues BGP update messages in case of a failure and acts as distributor for received update messages. The lightweight BGP engine can be seen as best case for BGP concerning the amount of update messages. In case of a failure the used BGP engine notifies every peer with exactly one BGP update message but filters duplicate updates. This behavior is rather unusual because BGP recalculates its routing table on the receipt of an update message and redistributes the new routes to its peers until a set of policy conforming routes is achieved.

The implemented FaSRo engine provides the functionality that was described in section III. In case of a link failure the FaSRo path is signaled from each failure detecting end point to the other.

The simulation intends to show:

- FaSRo provides a shorter convergence time
- Reduction of globally visible BGP update messages
- Routers are less stressed by update messages

A. Evaluation setup

To prove the FaSRo concept the topology depicted in Fig. 4 was used. Further this topology was used to validate that the implemented BGP and FaSRo engines work as intended.

For the topology the broken link was selected manually. The broken link has to fulfill the condition that an alternative path exists to substitute the broken link. Policies were not taken into account because it is assumed that ISPs, whose BGP routers are FaSRo capable, accept to transport traffic from other ISPs for a short period of time even if this would violate their policies.

TABLE I
Simulation Results

	10 ASes	
	FaSRo	BGP
Message ratio	22.2 %	100 %
Messages per total number of ASes	0.25	1.125
Ratio of convergence time	40 %	100 %
Min, Max length of FaSRo path	4, 5	–

B. Evaluation Results

The simulation has proven that the FaSRo concept works and provides an improvement to BGP. In the following, parameters from table I of the simulation are discussed:

The parameter *message ratio* correlates the total number of FaSRo messages to the total amount of BGP messages that occurred during the simulation. The first simulations have shown that FaSRo needs less than 25 % of the messages BGP needs to handle a link failure. The lightweight BGP engine—as already stated before—is an optimized version with respect to convergence time, which is directly correlated to the amount of update messages. So even in the best case BGP does not perform as good as FaSRo does.

The *messages per total number of ASes* correlates the total number of FaSRo and BGP messages to the number of ASes used in the simulation. This parameter offers an almost topology independent parameter describing the quality of the FaSRo process.

The *ratio of convergence time* correlates the convergence time FaSRo needs to establish a FaSRo path to the time BGP needs to propagate the link failure among the ASes. As mentioned before, a lightweight BGP engine is used that represents an optimal case concerning link failure propagation. As policies are out of scope during the first simulations BGP has converged as soon as all ASes are notified about the link failure. Providing absolute time values does not make sense because the internal router processing and the message transport time was not implemented in the simulation.

The *min and max length of FaSRo path* shows minimum and maximum the length of the substitution path—including the error detecting ASes—that is established to route around the link failure.

V. Conclusion and Outlook

The FaSRo process provides a mechanism to limit inter-domain network changes to only those peers that are necessarily affected by that change. Four main objectives can be achieved using FaSRo:

- Providing quickly an alternative path (if one exists) for sudden inter AS network disruptions
- Reduction of globally visible BGP update messages
- Thus reduction of routers' CPU and network load
- Opportunity for the Internet to converge to a coherent view

Those objectives are an essential premise with regard to the improvement of the current Internet. They aim at enabling the

network in order to provide stable routes as basis for future services with QoS guarantees.

First simulation results have shown qualitatively that the FaSRo approach satisfies its design goals.

Simulations with larger network scenarios, including real Internet topologies are to be addressed in future work. Furthermore, the simulations will be extended to provide a more realistic BGP behavior. It is also intended to extend the FaSRo approach with a kind of *whispering withdraw* so that peers that detect a network failure can ask their neighbors for alternative paths, if the failure detecting router as no alternative path already available.

VI. Appendix

Though FaSRo routers may have a different view than non-affected BGP routers, the following proof shows that the combination of FaSRo and BGP does not result in any loops, especially if FaSRo re-rerouting is active.

Notation: We model the Internet topology at the Autonomous System (AS) level as graph $G = (V, E)$ with $V := \{v | v \text{ is a node (AS)}\}$ and $E := \{(u, v) | u, v \in V\} \subseteq V \times V$. The *path* $p_{v_k}(d)$ denotes the sequence of ASes which are traversed towards destination network d as seen by AS v_k. Thus, if v_0 is the next AS hop and v_n the destination AS, $p_{v_k}(d)$ is defined as sequence of edges $(v_0, v_1), (v_1, v_2), \ldots, (v_{n-1}, v_n) = \langle v_0, v_1, \ldots, v_n \rangle$ with $(v_i, v_{i+1}) \in E$ for $i = 0, \ldots, n-1$.

Preconditions:

(i) Consistent BGP view (BGP had converged): $p_{v_k} = \langle v_{k+1}, \ldots, v_n \rangle = \langle v_{k+1}, p_{v_{k+1}}(d) \rangle$ for $k = 0, \ldots, n-1$. Thus, it follows that

$$\begin{aligned} p_{v_0}(d) &= \langle v_1, v_2, \ldots, v_n \rangle = \langle v_1, p_{v_1}(d) \rangle \\ &= \langle v_1, v_2, p_{v_2}(d) \rangle \\ p_{v_1}(d) &= \langle v_2, \ldots, v_n \rangle \\ p_{v_2}(d) &= \langle v_3, \ldots, v_n \rangle \\ \cdots &= \cdots \\ p_{v_{n-1}}(d) &= \langle v_n \rangle \\ p_{v_n}(d) &= \langle \rangle \end{aligned}$$

(ii) All BGP paths are loop-free. A path $p_{v_0}(d)$ is loop-free if $\forall v_i \in p_{v_k}(d) : v_i \neq v_j \ \forall i \neq j$

(iii) The original BGP path $p_{v_0}(d) = \langle v_1, v_2, \ldots, v_n \rangle$ is replaced by the path $p_{v_0}(d) = \langle \underbrace{u_1, \ldots, u_k}_{\substack{\text{FaSRo} \\ \text{loop-free path}}}, \underbrace{v_i, \ldots, v_n}_{\substack{\text{residual} \\ \text{path}}} \rangle$.

That means, without loss of generality, sub-path $\langle v_1, \ldots, v_{i-1} \rangle$ was substituted by the loop-free FaSRo path $\langle u_1, \ldots, u_k \rangle$, and, the residual BGP path $\langle v_i, \ldots, v_n \rangle$ remains unchanged.

Proof: Assume that at some AS v_l a node u_j of the FaSRo path is also present in the residual BGP path of $p_{v_l}(d)$ (i.e., a loop exists between BGP and FaSRo):

$$\begin{aligned} &\exists\, j \in \{1, \ldots, k\} : \exists\, l \in \{i, \ldots, n-1\} : u_j \in p_{v_l}(d) \\ &\Rightarrow\ p_{v_l}(d) = \langle v_{l+1}, \ldots, v_n \rangle = \langle v_{l+1}, \ldots, u_j, \ldots, v_n \rangle \\ &\overset{(i)}{\Rightarrow}\ p_{v_{l-1}}(d) = \langle v_l, p_{v_l}(d) \rangle = \langle v_l, v_{l+1}, \ldots, u_j, \ldots, v_n \rangle \\ &\Rightarrow\ p_{v_{l-2}}(d) = \langle v_{l-1}, p_{v_{l-1}}(d) \rangle = \langle v_{l-1}, v_l, \ldots, u_j, \ldots, v_n \rangle \\ &\Rightarrow\ \exists\, h : p_{v_h}(d) = \langle u_i, \ldots, u_k, v_l, \ldots, u_j, \ldots, v_n \rangle : u_i = u_j \\ &\overset{(ii)}{\Rightarrow}\ \text{contradiction to the loop-free property of BGP.} \blacksquare \end{aligned}$$

References

[1] Y. Rekhter and T. Li, "A Border Gateway Protocol 4 (BGP-4)," RFC 1771 (Standard), IETF, Mar. 1995.

[2] Y. Rekhter, T. Li, and S. Hares, "A Border Gateway Protocol 4 (BGP-4)," draft-ietf-idr-bgp4-19, IETF, Mar. 2003.

[3] G. Huston, "Scaling Inter-Domain Routing," *Internet Protocol Journal*, Dec. 2001. [Online]. Available: http://www.cisco.com/ipj

[4] ——, "Analyzing the Internet BGP Routing Table," *Internet Protocol Journal*, Mar. 2001. [Online]. Available: http://www.cisco.com/ipj

[5] J. Rexford, J. Wang, Z. Xiao, and Y. Zhang, "BGP routing stability of popular destinations," in *The Second Internet Measurement Workshop*. ACM SIGCOMM, Nov 2002, pp. 197–202.

[6] R. Mahajan, D. Wetherall, and T. Anderson, "Understanding BGP misconfiguration," in *Proceedings of ACM SIGCOMM 2002*, Pittsburgh, USA, Aug. 2002.

[7] C. Labovitz, G. R. Malan, and F. Jahanian, "Origins of Internet routing instability," in *Proceedings of IEEE INFOCOMM '99 Conference*, New York, New York, March 1999.

[8] R. Govindan and A. Reddy, "An analysis of inter-domain topology and route stability," in *Proceedings of IEEE INFOCOMM '97 Conference*, 1997.

[9] L. Subramanian, S. Agarwal, J. Rexford, and R. Katz, "Characterizing the Internet hierarchy from multiple vantage points," University of California," Paper, Aug. 2001.

[10] N. Taft, "The basics of BGP routing and its performance in today's Internet," RHDM (Reseaux Haut Debit et Multimedia). Corsica, France, May 2001.

[11] C. Labovitz, G. R. Malan, and F. Jahanian, "Internet routing instability," *IEEE/ACM Transactions on Networking*, vol. 6, no. 5, pp. 515–528, Oct. 1998. [Online]. Available: http://www.acm.org/pubs/citations/journals/ton/1998-6-5/p515-labovitz/

[12] D.-F. Chang, R. Govindan, and J. Heidemann, "An empirical study of router response to large BGP routing table load," in *The Second Internet Measurement Workshop*. ACM SIGCOMM, Nov 2002, pp. 203–208.

[13] L. Wang, X. Zhao, D. Pei, R. Bush, D. Massey, A. Mankin, S. Wu, and L. Zhang, "Observation and analysis of BGP behavior under stress," in *The Second Internet Measurement Workshop*. ACM SIGCOMM, Nov 2002, pp. 183–195.

[14] T. G. Griffin and B. J. Premore, "An Experimental Analysis of BGP Convergence Time," in *9th International Conference on Network Protocols, ICNP*, Nov. 2001.

[15] C. Labovitz, A. Ahuja, and A. Bose, "Delayed Internet routing convergence," in *Proceedings of ACM SIGCOMM 2000*, Stockholm, Sweden, August 2000.

[16] S. Sangli, Y. Rekhter, R. Fernando, J. Scudder, and E. Chen, "Graceful restart mechanism for BGP," draft-ietf-idr-restart-06.txt, Jul 2002.

[17] C. Villamizar, R. Chandra, and R. Govindan, "BGP Route Flap Damping," RFC 2439, IETF, Nov. 1999.

[18] G. Huston, "NOPEER community for BGP route scope control," draft-ietf-ptomaine-nopeer-02, IETF, Feb. 2003.

[19] R. Bush, "Route flap damping: harmful?" September 2002. [Online]. Available: http://psg.com/˜randy/020910.zmao-flap.pdf

[20] Z. M. Mao, R. Govind, G. Varghese, and R. H. Katz, "Route flap damping exacerbates Internet routing convergence," in *Proceedings of ACM SIGCOMM 2000*, Pittsburgh, USA, August 2002.

[21] S. Halabi and D. McPherson, *Internet Routing Architectures*. Cisco Press, 2000.

[22] A. Varga, "OMNeT++ Discrete Event Simulation System 2.2," May 2002. [Online]. Available: http://www.hit.bme.hu/phd/vargaa/omnetpp.htm