

Inter-Domain Routing  
Internet-Draft  
Intended status: Standards Track  
Expires: November 24, 2017

T. Kumar  
Ericsson  
May 23, 2017

**Split Horizon Considerations in Border Gateway Protocol (BGP) Fast  
ReRoute (FRR)  
draft-anup-idr-bgp-frr-splithorizon-00**

**Abstract**

This document describes the problems in certain types of BGP FRR deployments. It also suggests a solution in such scenarios. The described problem and solution are applicable to any distance vector routing protocol attempting FRR

**Requirements Language**

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC 2119](#) [[RFC2119](#)] .

**Status of This Memo**

This Internet-Draft is submitted in full conformance with the provisions of [BCP 78](#) and [BCP 79](#).

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on November 24, 2017.

**Copyright Notice**

Copyright (c) 2017 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to [BCP 78](#) and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of

publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

## Table of Contents

<a href="#">1.</a>	Introduction . . . . .	<a href="#">2</a>
<a href="#">2.</a>	Preamble . . . . .	<a href="#">2</a>
<a href="#">2.1.</a>	Sample Topology . . . . .	<a href="#">2</a>
<a href="#">2.2.</a>	Label exchange in FRR scenario . . . . .	<a href="#">3</a>
<a href="#">3.</a>	Problem . . . . .	<a href="#">3</a>
<a href="#">3.1.</a>	Summary of the Problem . . . . .	<a href="#">5</a>
<a href="#">4.</a>	Solution to the Problem . . . . .	<a href="#">5</a>
<a href="#">5.</a>	Acknowledgments . . . . .	<a href="#">6</a>
<a href="#">6.</a>	Security Considerations . . . . .	<a href="#">6</a>
<a href="#">7.</a>	IANA Considerations . . . . .	<a href="#">6</a>
<a href="#">8.</a>	Contributors . . . . .	<a href="#">6</a>
<a href="#">9.</a>	References . . . . .	<a href="#">6</a>
<a href="#">9.1.</a>	Normative References . . . . .	<a href="#">6</a>
<a href="#">9.2.</a>	Informative References . . . . .	<a href="#">7</a>
	Author's Address . . . . .	<a href="#">7</a>

## [1.](#) Introduction

BGP computes bestpaths, downloads them to RIB and advertises them to peer routers. When the bestpath is lost, BGP will recompute new bestpath, redownload to RIB and readvertise it to the peers. To be more resilient during bestpath loss, many vendor Implementations support a configuration called "frr" under bgp. When frr is enabled under bgp, bgp will compute a backup path for every bestpath and download the pair of paths to RIB, and advertise the pair of paths to peers (if allowed through another configuration called bgp add-paths advertise). In short, BGP FRR computes a backup path and gives them to RIB and peer routers..

## [2.](#) Preamble

The following sections describe the problems caused in certain BGP FRR deployments

### [2.1.](#) Sample Topology

Sample Topology



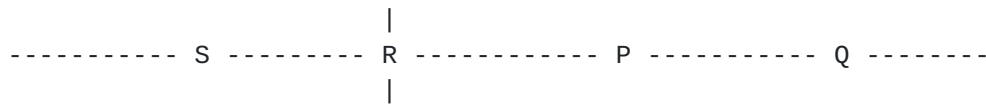


Figure1: Sample Topology

The path exchange on each router is as explained below:

- o Q - BGP router giving a path to peer router P for a destination D
- o P - BGP router giving the path (received from Q) to peer router R for destination D
- o S - BGP router giving a path to peer router R for destination D
- o R - BGP router receiving paths from routers S and P for destination D. Path from router S is chosen as bestpath. Path from router P is chosen as alternate path for destination D.

## 2.2. Label exchange in FRR scenario

The label exchange on each router is as explained below:

- o Suppose that BGP FRR is configured on R
- o Suppose that R is also configured for label exchange in BGP
- o R's bestpath is toward S, and backup path is towards P
- o R allocates a label Lr and advertises Lr to its peers
- o Since FRR is configured in bgp, internally R binds the label Lr to a pair of nexthops (S and P), from where bestpath and backup path were learnt respectively. This binding ensures that all label traffic carrying Lr and arriving at R, has primary nexthop as S, but also has a backup nexthop P.
- o R advertises its bestpath (S-R) and label Lr to all peers including P
- o It is important to note that, R's bestpath advertisement to P carries label Lr, while Lr is bound to nexthop pair S and P.

## 3. Problem

This section explains the problem when two BGP peers having FRR configuration, compute each other as backup nexthop. Suppose that BGP



FRR is configured on both R and P. Suppose that R and P are also configured for label exchange in BGP.

The sequence of path updates is as explained below:

- o Stage 1: R and P compute bestpath :
  - \* R receives path from S, and considers it as bestpath and advertises it to all peers along with label Lr1 (bound to bestpath).
  - \* P receives path from Q, and considers it as bestpath and advertises it to all peers along with label Lp1.
- o Stage 2: R receives backup path for first time :
  - \* R receives the above path from P, with a label Lp1, and considers it as backup path.
  - \* R allocates a label Lr2 (bound to pair of nexthops, S and P), and advertises its bestpath and label Lr2 to all peers including P.
  - \* Since Lr2 is bound to a pair of nexthops (S and P), it is called a double barrel label.
- o Stage 3: P receives backup path for first time :
  - \* P receives the above path from R, with label Lr2, and P considers it as backup path.
  - \* P allocates a label Lp2 (bound to pair of nexthops - R and Q) and advertises its bestpath and label Lp2 to all peers including R.
  - \* Since Lp2 is bound to a pair of nexthops (Q and R), it is called a double barrel label.
- o Stage 4: R receives backup path for second time :
  - \* R receives the 2nd advertisement from P with label Lp2
  - \* R sees that new path from P is different from old path received during (c) since label is different. So, R treats this as label change coming from P, and R will allocate a new label Lr3, and will advertise it to all peers, including P
- o Stage 5: P receives backup path for second time



- \* P receives the 3rd advertisement from R with label Lr3.
  - \* P sees it will be different from (e) since label is different. So, P treats this as label change coming from R, and P allocates a new label Lp4, and advertises it to all peers, including R.
- o Infinite loop :
- \* The above steps (Stage 2 to 5) repeat forever, and the two routers P and R will enter into a path advertisement loop.
  - \* It is important to note that stage 2 is the trigger for the problem, as it introduces the first label change

### **3.1. Summary of the Problem**

When two BGP routers, P and R are configured with bgp frr and bgp label exchange, and if they compute each other as backup nexthop, then they will enter into repeated label/path advertisement scenario, leading to route churn. This churn can be avoided by implementing split horizon rule at stage 2 in bgp frr, as explained in next section.

## **4. Solution to the Problem**

The advertisement churn explained in previous section occurs because of the way labels are bound and advertised in BGP when FRR is configured :

- o the allocated label (double barrel label) is bound to both bestpath and backup path, and is advertised to all peers, including the peer giving the backup path.
- o the double barrel label is unique to the 4-tuple : nexthop1, label from nexthop1, nexthop2, label from nexthop2
- o if the label from any of the nexthops changes, then a new double barrel label will be allocated and advertised. This happens in stage 2.
- o due to the above label change, the peer router sees a new label advertisement coming from the other peer, and it allocates a new label.
- o the above steps repeat forever



The above infinite loop triggered by stage 2 can be prevented in one of the following ways:

- o a. By not advertising the bestpath to the peer giving the backup nexthop in stage 2
- o b. By not advertising the double barrel label towards the backup nexthop in stage 2

Option (a) is against base bgp protocol, as the bestpath advertisement cannot be selectively blocked except by an administrative routemap. We prefer Option (b), i.e do not advertise double barrel label towards the backup nexthop. This is called as Split horizon rule in BGP FRR. Split horizon rule in BGP FRR can be defined as follows: "On a bgp router R having bgp frr configuration, when R computes bestpath towards S and backup path towards P, and allocates a double barrel label Lr for the pair of nexthops (S and P), R should not advertise the double barrel label Lr towards the backup nexthop P (since this label advertisement triggers the problem). Instead, R should advertise towards P the label L1 bound to the bestpath nexthop S. Further, R should advertise the double barrel label Lr to all peers, other than P." With the above split horizon rule, the routers R and P will be stable state after advertising the backup paths to each other only once.

## **5. Acknowledgments**

The authors would like to thank P. Muthu and team for their comments and review.

## **6. Security Considerations**

There are no additional security considerations than the base BGP RFC.

## **7. IANA Considerations**

This document has no actions for IANA.

## **8. Contributors**

## **9. References**

### **9.1. Normative References**



[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", [BCP 14](#), [RFC 2119](#), DOI 10.17487/RFC2119, March 1997, <<http://www.rfc-editor.org/info/rfc2119>>.

[RFC4271] Rekhter, Y., Ed., Li, T., Ed., and S. Hares, Ed., "A Border Gateway Protocol 4 (BGP-4)", [RFC 4271](#), DOI 10.17487/RFC4271, January 2006, <<http://www.rfc-editor.org/info/rfc4271>>.

## **9.2. Informative References**

[RFC3107] Rekhter, Y. and E. Rosen, "Carrying Label Information in BGP-4", [RFC 3107](#), DOI 10.17487/RFC3107, May 2001, <<http://www.rfc-editor.org/info/rfc3107>>.

### Author's Address

Anup Kumar T  
Ericsson India Pvt Ltd  
Ferns Icon, Doddanakkundi, Mahadevapura  
Bengaluru 560037  
India

Email: [anupkumar.t@ericsson.com](mailto:anupkumar.t@ericsson.com)

