Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: January 2, 2014

Pierre Francois
IMDEA Networks
Clarence Filsfils
Ahmed Bashandy
Stefano Previdi
Cisco Systems, Inc.
Bruno Decraene
Orange

July 1, 2013

Segment Routing Fast Reroute
draft-francois-sr-frr-00

Abstract

This document presents a Fast Reroute approach aimed at providing link protection of nodal and adjacency segments to the Segment Routing framework. This FRR behavior builds on proven IP-FRR concepts being LFAs, remote LFAs (RLFA), and remote LFAs with directed forwarding (DLFA). We describe their implementation using SR segments. We then analyze the benefits brought by Segment Routing to the scalability of such IP-FRR approaches.

Status of this Memo

Copyright Notice

Table of Contents

1. Introduction 3
2. Protection lists 4
2.1. LFA based protection list 4
2.2. RLFA based protection list 4
2.3. DLFA based protection list 5
3. Protecting segments 5
3.1. The active segment is a node segment 5
3.2. The active segment is an adjacency segment 6
3.2.1. Protecting [Adjacency, Adjacency] segment lists . . . 6
3.2.2. Protecting [Adjacency, Nodal] segment lists 6
4. SR FRR benefits in LDP environments 7
4.1. Eliminating Directed LDP Sessions 9
4.2. Guaranteed FRR coverage 9
5. References 10
Authors' Addresses 11

1. Introduction

Segment Routing aims at supporting services with tight SLA guarantees [1]. Acknowledging this fact, this document provides local repair mechanisms capable of restoring end-to-end connectivity in case of a sudden failure of a link. The FRR behavior builds on proven IP-FRR concepts; we leverage LFAs, remote LFAs (RLFA), and remote LFAs with directed forwarding (DLFA)[2].

In the SR context, not all flows are being routed along the shortest paths defined by the IGP, but also along explicit paths containing Adjacency Segments. We thus accommodate the IP-FRR behavior for SR Adjacency Segments.

Through the document, we will observe that performing FRR with SR has the following benefits:

- The simplicity properties of LFA FRR [3] are preserved.
- The capacity planning properties are preserved [3]. Unlike SDH and other FRR solutions, the repaired packet does not go back to the next-hop or next-next-hop but uses shortest-path forwarding from a much closer release point.
- The RLFA operation is simplified: dynamically established directed LDP sessions to the repair nodes are no longer required.
- The scalable support for DLFA provides guaranteed coverage for symmetric networks, i.e. networks configured with symmetric link metrics: the repair tunnel in a symmetric network can be encoded efficiently with only two segments. We will observe that only one segment is needed in most cases.

A future version of this document will analyze the protection upon node failure.

```
                                 L         ____
                            S----------F--{____}--D
                           _|_        ___________ /
                          {___}--R--{___________}
```

Figure 1: Link Protection

We use Figure 1 to illustrate the three objectives that have to be met when implementing SR FRR.

First, the protecting router S needs to find a detour path around the protected link. Intuitively, the Point of Local Repair (PLR) needs to find a node R (a repair node) that is capable of safely forwarding the traffic affected by the failure of the protected link L. We leverage the algorithms defined in the IP-FRR framework to achieve this first goal, as explained in Section 2.

Second, S must ensure proper forwarding behavior once the packet reaches the repair node R. We define the segment operations to be applied by the protecting node to ensure consistency with the forwarding state of the repair node in Section 3.

We will observe, in Section 4, that the MPLS instantiation of SR improves the scalability and operation of the FRR solution by not requiring multi-hop LDP sessions to distant repair nodes.

2. Protection lists

The protection list is the list of segments encoding a detour path from the protecting node S to the repair node R, avoiding the protected link L. In this section, we define how to encode the LFA, RLFA, and DLFA repair paths with protection lists. These protection lists contain at most two segments.

2.1. LFA based protection list

According to the LFA FRR approach, if a path to a destination D from a neighbor N of S does not contain S (i.e. N is a loop-free alternate of S for the failure of link S-F), then S can pre-install a repair forwarding information, in order to deviate the packet to N upon the failure of S-F.

In the case of LFA applicability, the SR protection list is thus empty. All what a protecting router S needs to do is to send the protected packet as is to its LFA neighbor N.

2.2. RLFA based protection list

If there is no such LFA neighbor, then S may be able to create a virtual LFA by using a tunnel to carry the packet to a point in the network that is not a direct neighbor of S, and from which the packet will be delivered to the destination without looping back to S. The Remote LFA proposal [4] calls such a tunnel a repair tunnel. The tail-end of this tunnel (R in figure 1) is called a "remote LFA" or a "PQ node". We refer to the RLFA document for the definitions of the P and Q sets.

In the case of RLFA applicability for the protection of a segment, the protection list is made of a nodal segment to the PQ node. It thus matches [nodal(PQ), ...]

2.3. DLFA based protection list

There are some cases where there is no remote LFA coverage for some links/destinations, due to topological properties in the neighborhood of the protecting node. If there is no such RLFA PQ node, we propose to use a Directed LFA (DLFA) repair tunnel to a Q node that is adjacent to the P space [5].

In the case of applicability of RLFA with directed forwarding (DLFA), the protection list is made of a nodal segment to the P node followed by an Adjacency segment to the Q node. It thus matches [nodal(P), Adj(P-->Q), ...]

In networks with symmetric IGP metrics (the metric of a link AB is the same as the metric of the reverse link BA), we can prove that either the P and the Q sets intersect or there is at least one P node that is adjacent to a Q node. Thanks to the DLFA extension, we thus have a guaranteed LFA-based FRR technique for any network with symmetric IGP metrics.

Future versions of the document will describe the solutions leveraging SR capabilities to provide guaranteed FRR applicability in any IGP topology.

3. Protecting segments

In this section, we explain how a protecting router S processes the active segment of a packet upon the failure of the primary adjacency along which the packet should be forwarded. The behavior depends on the type of active segment to be protected.

3.1. The active segment is a node segment

The definition of the protection of a nodal segment is a direct translation of IP-FRR behaviors into the SR terminology. That is, traffic for nodal segment D will be rerouted to a safe node R whose shortest paths for D do not contain the failed component.

As nodal segments semantics are known by all nodes of the domain, no specific signaling needs to be done to let R correctly process the detoured packet. A packet whose active segment matches [nodal(D),...], arriving at a protecting node S will leave S with a segment list matching [PS(R), nodal(D),...]. The actual value used

to encode nodal(D) is set by S based on the SRSB obtained from the IGP [1].

PS(R) is the computed Protection list to reach R, as discussed in section 2, and depends on the available type of protection: per-prefix LFA, Remote LFA or Directed LFA. The packet will follow the detour path defined by PS(R), and will finally reach R. When reaching R, the active segment of the packet is nodal(D), and the packet resumes its course along the original segment list.

3.2. The active segment is an adjacency segment

The operator may forbid protection of an adjacency segment by policy (?-Flag in [ISIS]/[OSPF]). For example, this is useful when the operator prefers an end-to-end protection mechanism triggered by the source of a multi-hop transport conduit.

We define hereafter the FRR behavior applied by S for any packet received with an active segment L for which protection was enabled. We distinguish the case where this active segment is followed by another adjacency segment from the case where it is followed by a nodal segment.

3.2.1. Protecting [Adjacency, Adjacency] segment lists

If the next segment in the list is an Adjacency Segment, then the packet has to be conveyed to F.

To do so, S applies a "NEXT" operation on Adj(L) and then two consecutive "PUSH" operations: first it pushes a nodal Segment for F, and then it pushes a protection list allowing to reach F while bypassing L.

Upon failure of L, a packet reaching S with a segment list matching [adj(L),adj(M),...] will thus leave S with a segment list matching [PS(F),nodal(F),adj(M)].

The protection list PS(F) will define the course of the packet from S to F, and F will resume the the course of the original segment list, receiving it with an active segment list matching [nodal(F), adj(M),...].

3.2.2. Protecting [Adjacency, Nodal] segment lists

If the next segment in the stack is a nodal segment, say for node T, the packet segment list matches [adj(L),nodal(T),...].

A first solution would consist in steering the packet back to F while

avoiding L, similarly to the previous case. To do so, S applies a "NEXT" operation on Adj(L) and then two consecutive "PUSH" operations: first it pushes a nodal Segment for F, and then it pushes a protection list allowing to reach F while bypassing L.

Upon failure of L, a packet reaching S with a segment list matching [adj(L),nodal(T),...] will thus leave S with a segment list matching [PS(F),nodal(F),nodal(T)].

Another solution is to not steer the packet back via F. In this case, S just needs to apply a "NEXT" operation on the Adjacency segment related to L, and push a protection segment list redirecting the traffic to a node R, capable of whose path to nodal segment T is not affected by the failure.

Upon failure of L, packets reaching S with a segment list matching [adj(L), nodal(T), ...], would leave S with a segment list matching [PS(R),nodal(T), ...].

4. SR FRR benefits in LDP environments

In this section, we describe the operational and scaling benefits of SR when used to implement RLFA and DLFA protection for LDP-based transport. We will also observe that a partial SR deployment, limited to the network region where the SR benefits are most desired, already provides the mentioned scaling benefits.

```
       X
       |
Y--A---B---E--Z
   |   |    \
   D---C--F--G
           30
```

Figure 2: Leveraging SR benefits for LDP-based traffic

In Figure 2, let us assume:

- All link costs are 10 except FG which is 30.
- All routers are LDP capable.
- X, Y and Z are PEs participating to an important service S.
- The operator requires 50msec link-based FRR for service S.
- A, B, C, D, E, F and G are SR capable.

- X, Y, Z are not SR capable. As part of a staged migration from LDP to SR, the operator deploys SR first in a sub-part of the network and then everywhere.

The operator would like to resolve the following issues:

- to protect the link BA along the shortest-path of the important flow XY, B requires an RLFA repair tunnel to D and hence a directed LDP session from B to D. The operator does not like these dynamically established multi-hop LDP sessions and would seek to eliminate them.
- there is no LFA/RLFA solution to protect the link BE along the shortest path of the important flow XZ. The operator wants a guaranteed link-based FRR solution.

The operator can meet these objectives by deploying SR only on A, B, C, D, E and F:

- The operator configures A, B, C, D, E, F and G with SRGB [100, 200] and respective node segments 101, 102, 103, 104, 105, 106 and 107.
- The operator configures D as an SR Mapping Server with the following policy mapping: (X, 201), (Y, 202), (Z, 203}.
- Each SR node automatically advertises local adjacency segment for its IGP adjacencies. Specifically, F advertises adjacency segment 9001 for its adjacency FG.

A, B, C, D, E, F and G keep their LDP capability and hence the flows XY and XZ are transported over end-to-end LDP LSP's.

For example, LDP at B installs the following MPLS dataplane entries:

- Incoming label: local LDB label bound by B for FEC Y
  - o Outgoing label: LDP label bound by A for FEC Y
  - o Outgoing nhop: A
- Incoming label: local LDB label bound by B for FEC Z
  - o Outgoing label: LDP label bound by E for FEC Z
  - o Outgoing nhop: E

The novelty comes from how the backup chains are computed for these LDP-based entries. While LDP labels are used for the primary nhop and outgoing labels, SR information is used for the FRR construction. In steady state, the traffic is transported over LDP LSP. In transient FRR state, the traffic is backed up thanks to the SR capabilities.

This helps meet the requirements of the operator:

- Eliminate directed LDP session
- Guaranteed FRR coverage

- Keep the traffic over LDP LSP in steady state
- Partial SR deployment only where needed

4.1. Eliminating Directed LDP Sessions

B's MPLS entry to Y becomes:

- Incoming label: local LDB label bound by B for FEC Y
    - o Outgoing label: LDP label bound by A for FEC Y
    - Backup outgoing label: SR node segment for Y {202}
    - o Outgoing nhop: A
    - Backup nhop: repair tunnel: node segment to D {104} with outgoing nhop: C

In steady-state, X sends its Y-destined traffic to B with a top label which is the LDP label bound by B for FEC Y. B swaps that top label for the LDP label bound by A for FEC Y and forwards to A. A pops the LDP label and forwards to Y.

Upon failure of the link BA, B swaps the incoming top-label with the node segment for Y (202) and sends the packet onto a repair tunnel to D (node segment 104). Thus, B sends the packet to C with the label stack {104, 202}. C pops the node segment 104 and forwards to D. D swaps 202 for 202 and forwards to A. A's nhop to Y is not SR capable and hence A swaps the incoming node segment 202 to the LDP label announced by its next-hop (in this case, implicit null).

After IGP convergence, B's MPLS entry to Y will become:

Incoming label: local LDB label bound by B for FEC Y
Outgoing label: LDP label bound by C for FEC Y
Outgoing nhop: C

And the traffic XY travels again over the LDP LSP.

The operator has eliminated its first problem: dynamically established directed LDP sessions are no longer required and the steady-state traffic is still transported over LDP. The SR deployment is confined to the area where these benefits were required.

4.2. Guaranteed FRR coverage

B's MPLS entry to Z becomes:

- Incoming label: local LDB label bound by B for FEC Z
    - Outgoing label: LDP label bound by E for FEC Z
    - o Backup outgoing label: SR node segment for Z {203}
    - Outgoing nhop: E

- o Backup nhop: repair tunnel to G: {106, 9001}
  - G is reachable from B via the combination of a node segment to F {106} and an adjacency segment FG {9001}
  - Note that {106, 107} would have equally work. Indeed, in many case, P's shortest path to Q is over the link PQ. The adjacency segment from P to Q is required only in very rare topologies where the shortest-path from P to Q is not via the link PQ.

In steady-state, X sends its Z-destined traffic to B with a top label which is the LDP label bound by B for FEC Z. B swaps that top label for the LDP label bound by E for FEC Z and forwards to E. E pops the LDP label and forwards to Z.

Upon failure of the link BE, B swaps the incoming top-label with the node segment for Z (203) and sends the packet onto a repair tunnel to G (node segment 106 followed by adjacency segment 9001). Thus, B sends the packet to C with the label stack {106, 9001, 203}. C pops the node segment 106 and forwards to F. F pops the adjacency segment 9001 and forwards to G. G swaps 203 for 203 and forwards to E. E's nhop to Z is not SR capable and hence E swaps the incoming node segment 203 for the LDP label announced by its next-hop (in this case, implicit null).

After IGP convergence, B's MPLS entry to Z will become:

- Incoming label: local LDB label bound by B for FEC Z
  - o Outgoing label: LDP label bound by C for FEC Z
  - o Outgoing nhop: C

And the traffic XZ travels again over the LDP LSP.

The operator has eliminated its second problem: guaranteed FRR coverage is provided. The steady-state traffic is still transported over LDP. The SR deployment is confined to the area where these benefits are required.

5. References

[1] Filsfils, C., Previdi, S., Bashandy, A., Decraene, B., Litkowski, S., Horneffer, M., Milojevic, I., Shakir, R., Ytti, S., Henderickx, W., Tantsura, J., and E. Crabbe, "Segment Routing Architecture", draft-filsfils-rtgwg-segment-routing-00 (work in progress), June 2013.

[2] Shand, M. and S. Bryant, "IP Fast Reroute Framework", RFC 5714, January 2010.

[3] Filsfils, C., Francois, P., Shand, M., Decraene, B., Uttaro, J., Leymann, N., and M. Horneffer, "Loop-Free Alternate (LFA) Applicability in Service Provider (SP) Networks", RFC 6571, June 2012.

[4] Bryant, S., Filsfils, C., Previdi, S., Shand, M., and S. Ning, "Remote LFA FRR", draft-ietf-rtgwg-remote-lfa-02 (work in progress), May 2013.

[5] Bryant, S., Filsfils, C., Previdi, S., and M. Shand, "IP Fast Reroute using tunnels", draft-bryant-ipfrr-tunnels-03 (work in progress), November 2007.

Authors' Addresses

Pierre Francois
IMDEA Networks
Leganes
ES

Email: pierre.francois@imdea.org

Clarence Filsfils
Cisco Systems, Inc.
Brussels
BE

Email: cfilsfil@cisco.com

Ahmed Bashandy
Cisco Systems, Inc.
San Jose
US

Email: bashandy@cisco.com

Stefano Previdi
Cisco Systems, Inc.
Rome
IT

Email: sprevidi@cisco.com

Bruno Decraene
Orange
Issy-les-Moulineaux
FR

Email: bruno.decraene@orange.com