Routing Area Working Group
Internet-Draft
Intended status: Standards Track
Expires: May 22, 2014

P. Sarkar, Ed.
H. Gredler
S. Hegde
H. Raghuveer
C. Bowers
Juniper Networks, Inc.
S. Litkowski
Orange

November 18, 2013

Remote-LFA Node Protection and Manageability
draft-psarkar-rtgwg-rlfa-node-protection-02

Abstract

The loop-free alternates computed following the current Remote-LFA [I-D.ietf-rtgwg-remote-lfa] specification gaurantees only link-protection. The resulting Remote-LFA nexthops (also called PQ-nodes), may not gaurantee node-protection for all destinations being protected by it.

This document describes procedures for determining if a given PQ-node provides node-protection for a specific destination or not. The document also shows how the same procedure can be utilised for collection of complete characteristics for alternate paths. Knowledge about the characteristics of all alternate path is precursory to apply operator defined policy for eliminating paths not fitting constraints.

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 RFC2119 [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 May 22, 2014.

Table of Contents

1. Introduction 3
2. Node Protection with Remote-LFA 3
2.1. The Problem 3
2.2. Few Additional Definitions 5
2.2.1. Link-Protecting Extended P-Space 5
2.2.2. Node-Protecting Extended P-Space 6
2.2.3. Q-Space 7
2.2.4. Link-Protecting PQ Space 7
2.2.5. Candidate Node-Protecting PQ Space 7
2.3. Computing Node-protecting R-LFA Path 7
2.3.1. Computing Candidate Node-protecting PQ-Nodes for Primary nexthops 8
2.3.2. Computing node-protecting paths from PQ-nodes to destinations 9
2.3.3. Limiting extra computational overhead 12
3. Manageabilty of Remote-LFA Alternate Paths 12
3.1. The Problem 12
3.2. The Solution 13
4. Acknowledgements 13
5. IANA Considerations 13
6. Security Considerations 13
7. References 13
7.1. Normative References 13
7.2. Informative References 14
Authors' Addresses 14

1. Introduction

The Remote-LFA [I-D.ietf-rtgwg-remote-lfa] specification provides loop-free alternates that gaurantees only link-protection. The resulting Remote-LFA alternate nexthops (also referred to as the PQ-nodes) may not provide node-protection for all destinations covered by the same, in case of failure of the primary nexthop node. Neither does the specification provide a means to determine the same.

Also, the LFA Manageability [I-D.ietf-rtgwg-lfa-manageability] document, requires a computing router to find all possible (including all possible Remote-LFA) alternate nexthops, collect the complete set of path characteristics for each alternate path, run a alternate-selection policy (configured by the operator), and find the best alternate path. This will require the Remote-LFA implementation to gather all the required path characteristics along each link on the entire Remote-LFA alternate path.

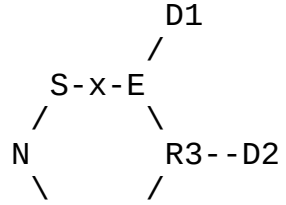
With current LFA [RFC5286] and Remote-LFA implementations, the forward SPF (and reverse SPF) is run on the computing router and its immediate 1-hop routers as the roots. While that enables computation of path attributes (e.g. SRLG, Admin-groups) for first alternate path segment from the computing router to the PQ-node, there is no means for the computing router to gather any path attributes for the path segment from the PQ-node to destination. Consecutively any policy-based selection of alternate paths will consider only the path attributes from the computing router up until the PQ-node.

This document describes a procedure for determining node-protection with Remote-LFA. The same procedure are also extended for collection of complete set of path attributes, enabling more accurate policy-based selection for alternate paths obtained with Remote-LFA.

2. Node Protection with Remote-LFA

2.1. The Problem

To better illustrate the problem and the solution proposed in this document the following topology diagram from the Remote-LFA [I-D.ietf-rtgwg-remote-lfa] draft is being re-used with slight modification.



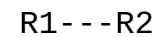


Figure 1: Topology 1

In the above topology, for all (non-ECMP) destinations reachable via the S-E link there is no standard LFA alternate. As per the Remote-LFA [I-D.ietf-rtgwg-remote-lfa] alternate specifications node R2 being the only PQ-node for the S-E link provides nexthop for all the above destinations. Table 1 below, shows all possible primary and Remote-LFA alternate paths for each destination.

Destination	Primary Path	PQ-node	Remote-LFA Backup Path
R3	S->E->R3	R2	S=>N=>R1=>R2->R3
E	S->E	R2	S=>N=>R1=>R2->R3->E
D1	S->E->D1	R2	S=>N=>R1=>R2->R3->E->D1
D2	S->E->R3->D2	R2	S=>N=>R1=>R2->R3->D2

Table 1: Remote-LFA backup paths via PQ-node R2

A closer look at Table 1 shows that, while the PQ-node R2 provides link-protection for all the destinations, it does not provide node-protection for destinations E and F. In the event of the node-failure on primary nexthop E, the alternate path from Remote-LFA nexthop R2 to E and D1 also becomes unavailable. So for a Remote-LFA nexthop to provide node-protection for a given destination, it is mandatory that, the shortest path from the given PQ-node to the given destination MUST not traverse the primary nexthop.

In another extension of the topology in Figure 1 let us consider an additional link between N and E.

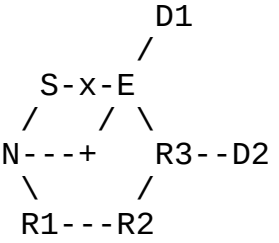


Figure 2: Topology 2

In the above topology, the S-E link is no more on any of the shortest paths from N to R3. Hence R3 is also included in both the Extended-P space and PQ space of E (w.r.t S-E link). Table 2 below, shows all possible primary and R-LFA alternate paths via PQ-node R3, for each

destination reachable through the S-E link in the above topology. The R-LFA alternate paths via PQ-node R2 remains same as in Table 1.

Destination	Primary Path	PQ-node	Remote-LFA Backup Path
R3	S->E->R3	R3	S=>N=>E=>R3
E	S->E	R3	S=>N=>E=>R3->E
D1	S->E->D1	R3	S=>N=>E=>R3->E->D1
D2	S->E->D1	R3	S=>N=>E=>R3->D2

Table 2: Remote-LFA backup paths via PQ-node R3

Again a closer look at Table 2 shows that, unlike Table 1, where the single PQ-node R2 provided node-protection, for destinations R3 and G, if we choose R3 as the R-LFA nexthop, it does not provide node-protection for R3 and D1 anymore. If S chooses R3 as the R-LFA nexthop, in the event of the node-failure on primary nexthop E, the alternate path from S to R-LFA nexthop R3 also becomes unavailable. So for a Remote-LFA nexthop to provide node-protection for a given destination, it is also mandatory that, the shortest path from S to the chosen PQ-node MUST not traverse the primary nexthop node.

2.2. Few Additional Definitions

This document adds and enhances the following definitions extending the ones mentioned in Remote-LFA [I-D.ietf-rtgwg-remote-lfa] draft.

2.2.1. Link-Protecting Extended P-Space

The Remote-LFA [I-D.ietf-rtgwg-remote-lfa] draft already defines this. The link-protecting extended P-space for a link S-E being protected is the set of routers that are reachable from one or more direct neighbors of S, except primary node E, without traversing the S-E link on any of the shortest path from the direct neighbor to the router. This MUST exclude any direct neighbor for which there is atleast one ECMP path from the direct neighbor traversing the link(S-E) being protected.

A node Y is in link-protecting extended P-space w.r.t to the link (S-E) being protected, if and only if, there exists atleast one direct neighbor of S, Ni, other than primary nexthop E, that satisfies the following condition.

$$D\_opt(Ni,Y) < D\_opt(Ni,S) + D\_opt(S,E) + D\_opt(E,Y)$$

Where,

```
D_opt(A,B) : Distance on most optimum path from A to B.
         E : The primary nexthop on shortest path from S
             to destination.
        Ni : A direct neighbor of S other than primary
             nexthop E.
         Y : The node being evaluated for link-protecting
             extended P-Space.
```

Figure 3: Link-Protecting Ext-P-Space Condition

2.2.2. Node-Protecting Extended P-Space

The node-protecting extended P-space for a primary nexthop node E being protected, is the set of routers that are reachable from one or more direct neighbors of S, except primary node E, without traversing the node E. This MUST exclude any direct neighbors for which there is atleast one ECMP path from the direct neighbor traversing the node E being protected.

A node Y is in node-protecting extended P-space w.r.t to the node E being protected, if and only if, there exists atleast one direct neighbor of S, Ni, other than primary nexthop E, that satisfies the following condition.

```
D_opt(Ni,Y) < D_opt(Ni,E) + D_opt(E,Y)
```

Where,

```
D_opt(A,B) : Distance on most optimum path from R1 to B.
         E : The primary nexthop on shortest path from S
             to destination.
        Ni : A direct neighbor of S other than primary
             nexthop E.
         Y : The node being evaluated for node-protecting
             extended P-Space.
```

Figure 4: Node-Protecting Ext-P-Space Condition

It must be noted that a node Y satisfying the condition in Figure 4 above only guarantees that the R-LFA alternate path segment from S via direct neighbor Ni to the PQ-node Y is not affected in the event of a node failure of E. It does not yet guarantee that the path segment from PQ-node Y to the destination is also unaffected by the same failure event.

2.2.3. Q-Space

The Remote-LFA [I-D.ietf-rtgwg-remote-lfa] draft already defines this. The Q-space for a link S-E being protected is the set of routers that can reach primary node E, without traversing the S-E link on any of the shortest path from the node Y to primary nexthop E. This MUST exclude any destination for which there is atleast one ECMP path from the node Y to the primary nexthop E traversing the link(S-E) being protected.

A node Y is in Q-space w.r.t to the link (S-E) being protected, if and only if, the following condition is satisfied.

D_opt(Y,E) < D_opt(S,E) + D_opt(Y,S)

Where,
  D_opt(A,B) : Distance on most optimum path from R1 to B.
          E : The primary nexthop on shortest path from S
               to destination.
          Y : The node being evaluated for Q-Space.

Figure 5: Q-Space Condition

2.2.4. Link-Protecting PQ Space

A node Y is in link-protecting PQ space w.r.t to the link (S-E) being protected, if and only if, Y is present in both link-protecting extended P-space and the Q-space for the link being protected.

2.2.5. Candidate Node-Protecting PQ Space

A node Y is in candidate node-protecting PQ space w.r.t to the node (E) being protected, if and only if, Y is present in both node-protecting extended P-space and the Q-space for the link being protected.

Again it must be noted that a node Y being in candidate node-protecting PQ-space does not guarantee that the R-LFA alternate path via the same, in entirety, is unaffected in the event of a node failure of primary nexthop node E. It only guarantees that the path segment from S to PQ-node Y is unaffected by the same failure event. The PQ-nodes in the candidate node-protecting PQ space may provide node protection for only a subset of destinations that are reachable through the corresponding primary link.

2.3. Computing Node-protecting R-LFA Path

The R-LFA alternate path through a given PQ-node to a given destination comprises of two path segments as follows.

1. Path segment from the computing router to the PQ-node (Remote-LFA alternate nexthop), and

2. Path segment from the PQ-node to the destination being protected.

So to ensure a R-LFA alternate path for a given destination provides node-protection we need to ensure that none of the above path segments are unaffected in the event of failure of the primary nexthop node. Sections Section 2.3.1 and Section 2.3.2 shows how this can be ensured.

2.3.1. Computing Candidate Node-protecting PQ-Nodes for Primary nexthops

To choose a node-protecting R-LFA nexthop for a destination R3, router S needs to consider a PQ-node from the candidate node-protecting PQ-space for the primary nexthop E on shortest path from S to R3. As mentioned in Section 2.2.2, to consider a PQ-node as candidate node-protecting PQ-node, there must be atleast one direct neighbor Ni of S, such that all shortest paths from Ni to the PQ-node does not traverse primary nexthop node E.

Implementations should run the inequality in Section 2.2.2 Figure 4 for all direct neighbor, other than primary nexthop node E, to determine whether a PQ-node Y is also a candidate node-protecting PQ-node. All of the metrics needed by this inequality would have been already collected from the forward SPFs rooted at each of direct neighbor S, computed as part of standard LFA [RFC5286] implementation. With reference to the topology in Figure 2, Table 3 below shows how the above condition can be used to determine the candidate node-protecting PQ-space for S-E link (primary nexthop E)

PQ-node (Y)	Direct Nbr (Ni)	D_opt (Ni,Y)	D_opt (Ni,E)	D_opt (E,Y)	Condition Met
R2	N	2 (N,R2)	1 (N,E)	2 (E,R2)	Yes
R3	N	2 (N,R3)	1 (N,E)	1 (E,R3)	No

Table 3: Node-protection evaluation for R-LFA repair tunnel to PQ-node

As seen in the above Table 3 , R3 does not meet the node-protecting extended-p-space inequality And so, while R2 is in candidate node-protecting PQ space, R3 is not.

Some SPF implementations may also produce a list of links and nodes traversed on the shortest path(s) from a given root to others. In such implementations, router S may have executed a forward SPF with each of it's direct neighbors as the SPF root, executed as part of the standard LFA [RFC5286] computations. So S may re-use the list of links and nodes collected from the same SPF computations, to decide whether a PQ-node Y is a candidate node-protecting PQ-node or not. A PQ-node Y shall be considered as a node-protecting, if and only if, there is atleast one direct neighbor of S, other than the primary nexthop E, for which, the primary nexthop node E does not exist on the list of nodes traversed on any of the shortest path(s) from the direct neighbor to the PQ-node. Table 4 below is an illustration of the mechanism with the topology in Figure 2.

PQ-node	Repair Tunnel Path(Repairing router to PQ-node)	Link-Protection	Node-Protection
R2	S->N->R1->R2	Yes	Yes
R2	S->E->R3->R2	No	No
R3	S->N->E->R3	Yes	No

Table 4: Protection of Remote-LFA tunnel to the PQ-node

As seen in the above Table 4 while R2 is candidate node-protecting Remote-LFA nexthop for R3 and G, it is not so for E and F, since the primary nexthop E is in the shortest path from R2 to E and F.

2.3.2. Computing node-protecting paths from PQ-nodes to destinations

Once a computing router finds all the candidate node-protecting PQ-nodes for a given directly attached primary link, it shall follow the procedure in proposed in this section, to choose one or more node-protecting R-LFA paths, for destinations reachable through the same primary link in the primary SPF graph.

To find a node-protecting R-LFA path for a given destination, the computing router needs to pick a subset of PQ-nodes from the candidate node-protecting PQ-space for the corresponding primary nexthop, such that all the path(s) from the PQ-node(s) to the given destination remain unaffected in the event of a node failure of

primary nexthop node. To ensure this, the computing router will need to ensure that, the primary nexthop node should not be on any of the shortest paths from the PQ-node to the given destination.

This document proposes an additional forward SPF computation for each of the PQ-nodes, to discover all shortest paths from the PQ-nodes to the destination. The additional forward SPF computation for each PQ-node, shall help determine, if a given primary nexthop node is on the shortest paths from the PQ-node to the given destination or not. To determine if a given PQ-node provides node-protecting alternate for a given destination, the primary nexthop node should not be on any of the shortest paths from the PQ-node to the given destination. After running the forward SPF on a PQ-node (from the node-protecting PQ-space) the computing router shall run the inequality in Figure 6 below. PQ-nodes that does not qualify the condition for a given destination, does not gaurantee node-protection for the path segment from the PQ-node to the given destination.

D_opt(Y,D) < D_opt(Y,E) + Distance_opt(E,D)

Where,
D_opt(A,B) : Distance on most optimum path from R1 to B.
D : The destination node.
E : The primary nexthop on shortest path from S to destination.
Y : The node-protecting PQ-node being evaluated

Figure 6: Node-Protecting Condition for PQ-node to Destination

All of the above metric costs except D_opt(Y, D), can be obtained with forward and reverse SPFs with E(the primary nexthop) as the root, run as part of the regular LFA and Remote-LFA implementation. The Distance_opt(Y, D) metric can only be determined by the additional forward SPF run with PQ-node Y as the root. With reference to the topology in Figure 2, Table 5 below shows how the above condition can be used to determine node-protection with node-protecting PQ-node R2.

Destination (D)	Primary-NH (E)	D_opt (Y, D)	D_opt (Y, E)	D_opt (E, D)	Condition Met
R3	E	1 (C,D)	2 (C,E)	1 (E,D)	Yes
E	E	2 (C,E)	2 (C,E)	0 (E,E)	No
D1	E	3 (C,F)	2 (C,E)	1 (E,F)	No

D2	E	2 (C,G)	2 (C,E)	1 (E,G)	Yes

Table 5: Node-protection evaluation for R-LFA path segment between PQ-node and destination

As seen in the above example above, R2 does not meet the node-protecting inequality for destination E, and F. And so, once again, while R2 is a node-protecting Remote-LFA nexthop for R3 and G, it is not so for E and F.

In SPF implementations that also produce a list of links and nodes traversed on the shortest path(s) from a given root to others, to determine whether a PQ-node provides node-protection for a given destination or not, the list of nodes computed from forward SPF run on the PQ-node, for the given destination, should be inspected. In case the list contains the primary nexthop node, the PQ-node does not provide node-protection. Else, the PQ-node guarantees node-protecting alternate for the given destination. Below is an illustration of the mechanism with candidate node-protecting PQ-node R2 in the topology in Figure 2.

Destination	Shortest Path (Repairing router to PQ-node)	Link-Protection	Node-Protection
R3	R2->R3	Yes	Yes
E	R2->R3->E	Yes	No
D1	R2->R3->E->D1	Yes	No
D2	R2->R3->D2	Yes	Yes

Table 6: Protection of Remote-LFA path between PQ-node and destination

As seen in the above example while R2 is candidate node-protecting R-LFA nexthop for R3 and G, it is not so for E and F, since the primary nexthop E is in the shortest path from R2 to E and F.

The procedure described in this document helps no more than to determine whether a given Remote-LFA alternate provides node-protection for a given destination or not. It does not find out any new Remote-LFA alternate nexthops, outside the ones already computed by standard Remote-LFA procedure. However, in case of availability of more than one PQ-node (Remote-LFA alternates) for a destination,

and node-protection is required for the given primary nexthop, this procedure will eliminate the PQ-nodes that do not provide node-protection and choose only the ones that does.

2.3.3. Limiting extra computational overhead

In addition to the extra reverse SPF computation, one per directly connected neighbor, suggested by the Remote-LFA [I-D.ietf-rtgwg-remote-lfa] draft, this document proposes a forward SPF per PQ-node discovered in the network. Since the average number of PQ-nodes found in any network is considerably more than the number of direct neighbors of the computing router, the proposal of running one forward SPF per PQ-node may add considerably to the overall SPF computation time.

To limit the computational overhead of the approach proposed, this document proposes that implementations MUST choose a subset from the entire set of PQ-nodes computed in the network, with a finite limit on the number of PQ-nodes in the subset. Implementations MUST choose a default value for this limit and may provide user with a configuration knob to override the default limit. Implementations MUST also evaluate some default preference criteria while considering a PQ-node in this subset. Finally, implementations MAY also allow user to override the default preference criteria, by providing a policy configuration for the same.

A suggested default criteria for PQ-node selection will be to put a score on each PQ-node, proportional to the number of primary interfaces and remote destination routers being protected by it, and then pick PQ-nodes based on this score. A more appropriate heuristsics can be devised, based on in-depth study of coverage provided by R-LFA, in the networks where they are mostly deployed. The same can then be used for PQ-node selection.

Once a subset of PQ-nodes is found, computing router shall run a forward SPF on each of the PQ-nodes in the subset to continue with procedures proposed in section Section 2.3.2.

3. Manageabilty of Remote-LFA Alternate Paths

3.1. The Problem

With the regular Remote-LFA [I-D.ietf-rtgwg-remote-lfa] functionality the computing router may compute more than one PQ-node as usable Remote-LFA alternate nexthops. Additionally an alternate selection policy may be configured to enable the network operator to choose one of them as the most appropriate Remote-LFA alternate. For such policy-based alternate selection to run, all the relevant path

characteristics for each the alternate paths (one through each of the PQ-nodes), needs to be collected. As mentioned befor in section Section 2.3 the R-LFA alternate path through a given PQ-node to a given destination comprises of two path segments.

The first path segment (i.e. from the computing router to the PQ-node) can be calculated from the regular forward SPF done as part of standard and remote LFA computations. However without the mechanism proposed in section Section 2.3.2 of this document, there is no way to determine the path characteristics for the second path segment (i.e from the PQ-node to the destination). In the absence of the path characteristics for the second path segment, two Remote-LFA alternate path may be equally preferred based on the first path segments characteristics only, although the second path segment attributes may be different.

3.2. The Solution

The additional forward SPF computation proposed in section Section 2.3.2 document shall also collect links, nodes and path characteristics along the second path segment. This shall enable collection of complete path characteristics for a given Remote-LFA alternate path to a given destination. The complete alternate path characteristics shall then facilitate more accurate alternate path selection while running the alternate selection policy.

4. Acknowledgements

Many thanks to Bruno Decraene for his useful comments.

5. IANA Considerations

N/A. - No protocol changes are proposed in this document.

6. Security Considerations

This document does not introduce any change in any of the protocol specifications. It simply proposes to run an extra SPF rooted on each PQ-node discovered in the whole network.

7. References

7.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

7.2. Informative References

[I-D.ietf-rtgwg-lfa-manageability]
Litkowski, S., Decraene, B., Filsfils, C., and K. Raza, "Operational management of Loop Free Alternates", draft-ietf-rtgwg-lfa-manageability-00 (work in progress), May 2013.

[I-D.ietf-rtgwg-remote-lfa]
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.

[I-D.litkowski-rtgwg-node-protect-remote-lfa]
Litkowski, S., "Node protecting remote LFA", draft-litkowski-rtgwg-node-protect-remote-lfa-00 (work in progress), April 2013.

[RFC5286] Atlas, A. and A. Zinin, "Basic Specification for IP Fast Reroute: Loop-Free Alternates", RFC 5286, September 2008.

Authors' Addresses

Pushpasis Sarkar (editor)
Juniper Networks, Inc.
Electra, Exora Business Park
Bangalore, KA 560103
India

Email: psarkar@juniper.net

Hannes Gredler
Juniper Networks, Inc.
1194 N. Mathilda Ave.
Sunnyvale, CA 94089
US

Email: hannes@juniper.net

Shraddha Hegde
Juniper Networks, Inc.
Electra, Exora Business Park
Bangalore, KA 560103
India

Email: shraddha@juniper.net

Harish Raghuveer
Juniper Networks, Inc.
Electra, Exora Business Park
Bangalore, KA 560103
India

Email: hraghuveer@juniper.net

Chris Bowers
Juniper Networks, Inc.
1194 N. Mathilda Ave.
Sunnyvale, CA 94089
US

Email: cbowers@juniper.net

Stephane Litkowski
Orange

Email: stephane.litkowski@orange.com