Routing Area Working Group   P. Sarkar, Ed.
Internet-Draft   Individual Contributor
Intended status: Standards Track   S. Hegde
Expires: June 26, 2017   C. Bowers
                                                   Juniper Networks, Inc.
                                                              H. Gredler
                                                           RtBrick, Inc.
                                                            S. Litkowski
                                                                  Orange
                                                       December 23, 2016

            Remote-LFA Node Protection and Manageability
             draft-ietf-rtgwg-rlfa-node-protection-09

Abstract

   The loop-free alternates computed following the current Remote-LFA
   specification guarantees only link-protection. The resulting Remote-
   LFA nexthops (also called PQ-nodes), may not guarantee node-
   protection for all destinations being protected by it.

   This document describes an extension to the Remote Loop-Free based IP
   fast reroute mechanisms described in [RFC7490], that 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 June 26, 2017.

Copyright Notice

Table of Contents

1. Introduction

   The Remote-LFA [RFC7490] specification provides loop-free alternates
   that guarantee 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 [RFC7916] 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 an 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 with 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.
   Consequently 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 is also extended for collection
   of a complete set of path attributes, enabling more accurate policy-
   based selection for alternate paths obtained with Remote-LFA.

2. Node Protection with Remote-LFA

   Node-protection is required to provide protection of traffic on a
   given forwarding node, against the failure of the first-hop node on
   the primary forwarding path. Such protection becomes more critical
   in the absence of mechanisms like non-stop-routing in the network.

Certain operators refrain from deploying non-stop-routing in their network, due to the required complex state synchronization between redundant control plane hardwares it requires, and the significant additional performance complexities it hence introduces. In such cases node-protection is essential to guarantee un-interrupted flow of traffic, even in the case of an entire forwarding node going down.

The following sections discuss the node-protection problem in the context of Remote-LFA and propose a solution.

2.1. The Problem

To better illustrate the problem and the solution proposed in this document the following topology diagram from the Remote-LFA [RFC7490] draft is being re-used with slight modification.

```
                        D1
                        /
                     S-x-E
                     /     \
                    N       R3--D2
                     \     /
                     R1---R2
```
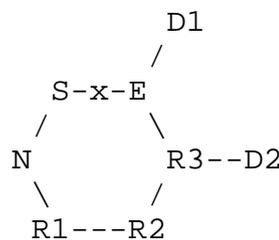
Figure 1: Topology 1

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 [RFC7490] 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 D1. 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 with the same cost as the other
links.

```
                               D1
                              /
                           S-x-E
                          /   / \
                       N---+    R3--D2
                         \      /
                         R1---R2
```
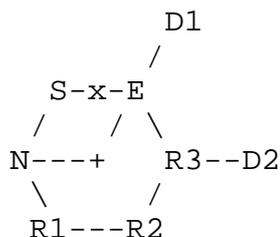
Figure 2: Topology 2

In the above topology, the S-E link is no more on any of the shortest
paths from N to R3, E and D1. Hence R3, E and D1 are also included
in both the Extended-P space and Q 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->R3->D2 | 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
D2, if we choose R3 as the R-LFA nexthop, it does not provide node-
protection for R3 and D2 anymore. If S chooses R3 as the R-LFA
nexthop, in the event of the node-failure on primary nexthop E, on
the alternate path from S to R-LFA nexthop R3, one of parallel ECMP
path between N and 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. Additional Definitions

   This document adds and enhances the following definitions extending
   the ones mentioned in Remote-LFA [RFC7490] specification.

2.2.1. Link-Protecting Extended P-Space

   The Remote-LFA [RFC7490] specification 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 at
   least one ECMP path from the direct neighbor traversing the link(S-E)
   being protected.

   For a cost-based definition for Link-protecting Extended P-Space
   refer to Section 2.2.6.1.

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 at least one ECMP path from the direct neighbor traversing the
   node E being protected.

   For a cost-based definition for Node-protecting Extended P-Space
   refer to Section 2.2.6.2.

2.2.3. Q-Space

   The Remote-LFA [RFC7490] 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 at least one ECMP path
   from the node Y to the primary nexthop E traversing the link(S-E)
   being protected.

   For a cost-based definition for Q-Space refer to Section 2.2.6.3.

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.

   Please note, 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.

2.2.6. Cost-Based Definitions

   This section provides cost-based definitions for some of the terms
   introduced in Section 2.2 of this document.

2.2.6.1. Link-Protecting Extended P-Space

   Please refer to Section 2.2.1 for a formal definition for Link-
   protecting Extended P-Space.

   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 at least one
   direct neighbor of S, Ni, other than primary nexthop E, that
   satisfies the following condition.

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

   Where,
      $D\_opt(A,B)$: Distance on most optimum path from A to B.
            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

2.2.6.2. Node-Protecting Extended P-Space

   Please refer to Section 2.2.2 for a formal definition for Node-
   protecting Extended P-Space.

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 at least 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 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 node-protecting
                extended P-Space.

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

Please note, 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 node Y is not affected in the event of a node failure of E. It does not yet guarantee that the path segment from node Y to the destination is also unaffected by the same failure event.

2.2.6.3. Q-Space

   Please refer to Section 2.2.3 for a formal definition for Q-Space.

   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 A 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

2.3. Computing Node-protecting R-LFA Path

   The R-LFA alternate path through a given PQ-node to a given
   destination is comprised 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 affected 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 at least 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 neighbors, other than primary nexthop node E, to
   determine whether a node Y is 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).

Candidate 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 its 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 node Y is a candidate node-protecting PQ-node or not. A node Y shall be considered as a node-protecting PQ-node, if and only if, there is at least 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.

Candidate 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 D2, it is not so for E and D1, since the primary nexthop E is in the shortest path from R2 to E and D1.

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 as 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 the
   primary nexthop node. To determine wether a given PQ-node belongs to
   such a subset of PQ-nodes, the computing router MUST ensure that none
   of the primary nexthop node are found 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. This will 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 candidate node-
   protecting PQ-node provides node-protecting alternate for a given
   destination, or not, all the shortest paths from the PQ-node to the
   given destination has to be inspected, to check if the primary
   nexthop node is found on any of these shortest paths. To compute all
   the shortest paths from a candidate node-protecting PQ-node to one
   (or more) destination, the computing router MUST run the forward SPF
   on the candidate node-protecting PQ-node. Soon after running the
   forward SPF, the computer router SHOULD run the inequality in
   Figure 6 below, once for each destination. A PQ-node that does not
   qualify the condition for a given destination, does not guarantee
   node-protection for the path segment from the PQ-node to the specific
   destination.

   $D\_opt(Y,D) < D\_opt(Y,E) + Distance\_opt(E,D)$

   Where,
      $D\_opt(A,B)$: Distance on most optimum path from A 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 (R2,R3)	2 (R2,E)	1 (E,R3)	Yes
E	E	2 (R2,E)	2 (R2,E)	0 (E,E)	No
D1	E	3 (R2,D1)	2 (R2,E)	1 (E,D1)	No
D2	E	2 (R2,D2)	2 (R2,E)	1 (E,D2)	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 D1. And so, once again,
while R2 is a node-protecting Remote-LFA nexthop for R3 and D2, it is
not so for E and D1.

In SPF implementations that also produce a list of links and nodes
traversed on the shortest path(s) from a given root to others, the
inequality in Figure 6 above need not be evaluated. Instead, 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  | Link-Protection | Node-Protection |
|            | (Repairing      |                |                |
|            | router to PQ-   |                |                |
|            | node)           |                |                |
+------------+-----------------+----------------+----------------+
| 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 D2, it is not so for E and D1, since the
primary nexthop E is in the shortest path from R2 to E and D1.

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. Computing Node-Protecting R-LFA Paths for Destinations with ECMP
        primary nexthop nodes

In certain scenarios, when one or more destinations maybe reachable
via multiple ECMP (equal-cost-multi-path) nexthop nodes, and only
link-protection is required, there is no need to compute any
alternate paths for such destinations. In the event of failure of
one of the nexthop links, the remaining primary nexthops shall always
provide link-protection. However, if node-protection is required,
the rest of the primary nexthops may not gaurantee node-protection.
Figure 7 below shows one such example topology.

```
                              D1
                    2       /
                   S---x---E1
                  / \      / \
                 /   x    /   \
                /     \  /     \
               N-------E2      R3--D2
                \    2        /
                 \          /
                  \        /
                  R1-------R2
                      2
```
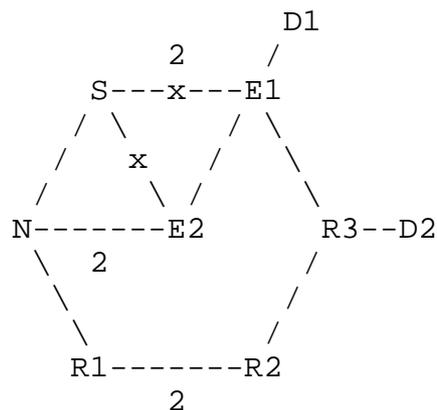
Primary Nexthops:
  Destination D1 = [{ S-E1, E1}, {S-E2, E2}]
  Destination D2 = [{ S-E1, E1}, {S-E2, E2}]

       Figure 7: Toplogy with multiple ECMP primary nexthops

In the above example topology, costs of all links are 1, except the
following links:

    Link: S-E1, Cost: 2

    Link: N-E2: Cost: 2

    Link: R1-R2: Cost: 2

In the above topology, on computing router S, destinations D1 and D2
are reachable via two ECMP nexthop nodes E1 and E2. However the
primary paths via nexthop node E2 also traverses via the nexthop node
E1. So in the event of node failure of nexthop node E1, both primary
paths (via E1 and E2) becomes unavailable. Hence if node-protection
is desired for destinations D1 and D2, alternate paths that does not
traverse any of the primary nexthop nodes E1 and E2, need to be
computed. In the above topology the only alternate neighbor N does
not provide such a LFA alternate path. Hence one (or more) R-LFA
node-proecting alternate paths for destinations D1 and D2, needs to
be computed.

In the above topology, following are the link-protecting PQ-nodes.

    Primary Nexthop: E1, Link-Protecting PQ-Node: { R2 }

    Primary Nexthop: E2, Link-Protecting PQ-Node: { R2 }

To find one (or more) node-protecting R-LFA paths for destinations D1
and D2, one (or more) node-protecting PQ-node(s) needs to be

determined first. Inequalities specified in Section 2.2.6.2 and
Section 2.2.6.3 can be evaluated to compute the node-protecting PQ-
space for each of the nexthop nodes E1 and E2, as shown in Table 7
below. To select a PQ-node as node-protecting PQ-node for a
destination with multiple primary nexthop nodes, the PQ-node MUST
satisfy the inequality for all primary nexthop nodes. Any PQ-node
which is NOT node-protecting PQ-node for all the primary nexthop
nodes, MUST NOT be chosen as the node-protecting PQ-node for
destination.

Primary Nexthop (E)	Candidate PQ-node (Y)	Direct Nbr (Ni)	D_opt (Ni,Y)	D_opt (Ni,E)	D_opt (E,Y)	Condition Met
E1	R2	N	3 (N,R2)	3 (N,E1)	2 (E1,R2)	Yes
E2	R2	N	3 (N,R2)	2 (N,E2)	3 (E2,R2)	Yes

Table 7: Computing Node-protected PQ-nodes for nexthop E1 and E2

In SPF implementations that also produce a list of links and nodes
traversed on the shortest path(s) from a given root to others, the
tunnel-repair paths from the computing router to candidate PQ-node
can be examined to ensure that none of the primary nexthop nodes is
traversed. PQ-nodes that provide one (or more) Tunnel-repair
paths(s) that does not traverse any of the primary nexthop nodes, are
to be considered as node-protecting PQ-nodes. Table 8 below shows
the possible tunnel-repair paths tp PQ-node R2.

Primary-NH (E)	PQ-Node (Y)	Tunnel-Repair Paths	Exclude All Primary-NH
E1, E2	R2	S==>N==>R1==>R2	Yes

Table 8: Tunnel-Repair paths to PQ-node R2

From Table 7 and Table 8, in the above example, R2 being node-
protecting PQ-node for both primary nexthops E1 and E2, should be
chosen as the node-protecting PQ-node for destinations D1 and D2 that
are both reachable via primary nexthop nodes E1 and E2.

Next, to find a node-protecting R-LFA path from node-protecting PQ-
node to destinations D1 and D2, inequalities specified in Figure 6
should be evaluated, to ensure if R2 provides a node-protecting R-LFA
path for each of these destinations, as shown below in Table 9. For
a R-LFA path to qualify as node-protecting R-LFA path for a
destination with multiple ECMP primary nexthop nodes, the R-LFA path
from the PQ-node to the destination MUST satisfy the inequality for
all primary nexthop nodes.

Destinat ion (D)	Primary- NH (E)	PQ- Node (Y)	D_opt (Y, D)	D_opt (Y, E)	D_opt (E, D)	Conditio n Met
D1	E1	R2	3 (R2, D1)	2 (R2, E1)	1 (E1, D1)	No
D1	E2	R2	3 (R2, D1)	3 (R2, E2)	2 (E2, D1)	Yes
D2	E1	R2	2 (R2, D2)	2 (R2, E1)	2 (E1, D2)	Yes
D2	E2	R2	2 (R2, D2)	2 (R2, E2)	3 (E2, D2)	Yes

Table 9: Finding node-protecting R-LFA path for destinations D1 and
D2

In SPF implementations that also produce a list of links and nodes
traversed on the shortest path(s) from a given root to others, the
R-LFA paths via node-protecting PQ-node to final destination can be
examined to ensure that none of the primary nexthop nodes is
traversed. R-LFA path(s) that does not traverse any of the primary
nexthop nodes, gaurantees node-protection in the event of failure of
any of the primary nexthop nodes. Table 10 below shows the possible
R-LFA-paths for destinations D1 and D2 via the node-protecting PQ-
node R2.

Destination (D)	Primary-NH (E)	PQ-Node (Y)	R-LFA Paths	Exclude All Primary-NH
D1	E1, E2	R2	S==>N==>R1==>R2 -->R3-->E1-->D1	No
D2	E1, E2	R2	S==>N==>R1==>R2 -->R3-->D2	Yes

Table 10: R-LFA paths for destinations D1 and D2

From Table 9 and Table 10, in the above example above, the R-LFA path from R2 does not meet the node-protecting inequality for destination D1, while it does meet the same inequality for destination D2. And so, while R2 provides node-protecting R-LFA alternate for D2, it fails to provide node-protection for destination D1. Finally, while it is possible to get a node-protecting R-LFA path for D2, no such node-protecting R-LFA path can be found for D1.

2.3.4. Limiting extra computational overhead

In addition to the extra reverse SPF computations suggested by the Remote-LFA [RFC7490] draft (one reverse SPF for each of the directly connected neighbors), this document proposes a forward SPF computations for each 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.

This document proposes that implementations SHOULD use a default preference criteria for PQ-node selection which will put a score on each PQ-node, proportional to the number of primary interfaces for which it provides coverage, its distance from the computing router,

and its router-id (or system-id in case of IS-IS). PQ-nodes that
cover more primary interfaces SHOULD be preferred over PQ-nodes that
cover fewer primary interfaces. When two or more PQ-nodes cover the
same number of primary interfaces, PQ-nodes which are closer (based
on metric) to the computing router SHOULD be preferred over PQ-nodes
farther away from it. For PQ-nodes that cover the same number of
primary interfaces and are the same distance from the the computing
router, the PQ-node with smaller router-id (or system-id in case of
IS-IS) SHOULD be preferred.

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 [RFC7490] 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 before in section Section 2.3 the R-LFA
alternate path through a given PQ-node to a given destination is
comprised 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 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.

As already specified in Section 2.3.4 to limit the computational overhead of the approach proposed, forward SPF computations MUST be run on a selected 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. The detailed suggestion on how to select this subset is specified in the same section. While this limits the number of possible alternate paths provided to the alternate-selection policy, this is needed keep the computational complexity within affordable limits. However if the alternate-selection policy is very restrictive this may leave few destinations in the entire toplogy without protection. Yet this limitation provides a necessary tradeoff between extensive coverage and immense computational overhead.

4. Acknowledgements

Many thanks to Bruno Decraene for providing his useful comments. We would also like to thank Uma Chunduri for reviewing this document and providing valuable feedback. Also, many thanks to Harish Raghuveer for his review and comments on the initial versions of this document.

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,
              DOI 10.17487/RFC2119, March 1997,
              <http://www.rfc-editor.org/info/rfc2119>.

   [RFC5286] Atlas, A., Ed. and A. Zinin, Ed., "Basic Specification for
              IP Fast Reroute: Loop-Free Alternates", RFC 5286,
              DOI 10.17487/RFC5286, September 2008,
              <http://www.rfc-editor.org/info/rfc5286>.

   [RFC7490] Bryant, S., Filsfils, C., Previdi, S., Shand, M., and N.
              So, "Remote Loop-Free Alternate (LFA) Fast Reroute (FRR)",
              RFC 7490, DOI 10.17487/RFC7490, April 2015,
              <http://www.rfc-editor.org/info/rfc7490>.

7.2. Informative References

   [RFC7916] Litkowski, S., Ed., Decraene, B., Filsfils, C., Raza, K.,
              Horneffer, M., and P. Sarkar, "Operational Management of
              Loop-Free Alternates", RFC 7916, DOI 10.17487/RFC7916,
              July 2016, <http://www.rfc-editor.org/info/rfc7916>.

Authors' Addresses

   Pushpasis Sarkar (editor)
   Individual Contributor

   Email: pushpasis.ietf@gmail.com

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

   Email: shraddha@juniper.net

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

   Email: cbowers@juniper.net

   Hannes Gredler
   RtBrick, Inc.

   Email: hannes@rtbrick.com

   Stephane Litkowski
   Orange

   Email: stephane.litkowski@orange.com