Working Group
Internet-Draft
Intended status: Informational
Expires: March 12, 2016

U. Chunduri
J. Tantsura
Ericsson Inc.
C. Bowers
Juniper Networks

September 9, 2015

Extended procedures and considerations for evaluating Loop-Free Alternates

draft-chunduri-rtgwg-lfa-extended-procedures-03

Abstract

This document provide few clarifications and extended procedures to IP Fast Reroute using Loop-Free Alternates as defined in RFC 5286.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

Loop Free Alternatives (LFAs) as defined in [RFC5286] have been widely deployed, and the operational and manageability considerations are described in great detail in [I-D.ietf-rtgwg-lfa-manageability].

This document intends to provide clarifications, additional considerations to [RFC5286], to address a few coverage and operational observations. These observations are in the area of handling Muti-homed prefixes (MHPs), IS-IS attach (ATT) bit in L1 area, links provisioned with MAX_METRIC for traffic engineering (TE) purposes and in the area of Multi Topology (MT) IGP deployments. All these are elaborated in detail in Section 2.

1.1. 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].

1.2. Acronyms

AF - Address Family

ATT - IS-IS Attach Bit

ECMP - Equal Cost Multi Path

IGP - Interior Gateway Protocol

IS-IS - Intermediate System to Intermediate System

OSPF - Open Shortest Path First

MHP - Multi-homed Prefix

MT - Multi Topology

SPF - Shortest Path First PDU

2. LFA Extended Procedures

This section explains the additional considerations in various aspects as listed below to the base LFA specification [RFC5286].

2.1. Multi Homed Prefixes

LFA base specification [RFC5286] Section 6.1 recommends that a router compute the alternate next-hop for an IGP multi-homed prefix by considering alternate paths via all routers that have announced that prefix. However, it also allows for the router to simplify the multi-homed prefix calculation by assuming that the MHP is solely attached to the router that was its pre-failure optimal point of attachment, at the expense of potentially lower coverage. If an implementation chooses to simplify the multi-homed prefix calculation by assuming that the MHP is solely attached to the router that was its pre-failure optimal point of attachment, the procedure described in this memo can potentially improve coverage for equal cost multi path (ECMP) MHPs without incurring extra computational cost.

The approach as specified in [RFC5286] Section 6.1 last paragraph, is to simplify the MHP is solely attached to the router that was its pre-failure optimal point of attachment. While this is very scalable approach and simplifies computation, as [RFC5286] notes this may result in little less coverage.

This memo improves the above approach to provide loop-free alternatives without any additional cost for equal cost multi path MHPs as described through the below example network. The approach specified here MAY also applicable for handling default routes as explained in Section 2.1.1.

```
       5   +---+  8   +---+  5  +---+
     +-----| S |------| A |-----| B |
     |     +---+      +---+     +---+
     |       |                    |
     |     5 |                  5 |
     |       |                    |
   +---+ 5 +---+   4 +---+  1    +---+
   | C |---| E |-----| M |-------| F |
   +---+   +---+     +---+       +---+
             |   10           5    |
             +-----------p---------+
```

Figure 1: MHP with same ECMP Next-hop

In the above network a prefix p, is advertised from both Node E and Node F. With simplified approach taken as specified in [RFC5286] Section 6.1, prefix p will get only link protection LFA through the neighbor C while a node protection path is available through neighbor A. In this scenario, E and F both are pre-failure optimal points of attachment and share the same primary next-hop. Hence, an implementation MAY compare the kind of protection A provides to F (link-and-node protection) with the kind of protection C provides to E (link protection) and inherit the better alternative to prefix p and here it is A.

However, in the below network prefix p has an ECMP through both node E and node F with cost 20. Though it has 2 pre-failure optimal points of attachment, the primary next-hop to each pre-failure optimal point of attachment is different. In this case, prefix p shall inherit corresponding LFA to each primary next-hop calculated for the router advertising the same respectively (node E’s and node F’s LFA).

```
           +---+         3      +---+
           | S |----------------| B |
           +---+                +---+
             |                    |
          10 |                  1 |
             |                    |
           +---+        6        +---+
           | E |-----------------| F |
           +---+                 +---+
             |   10          16    |
             +-----------p---------+
```

Figure 2: MHP with different ECMP Next-hops

In summary, if there are multiple pre-failure points of attachment for a MHP and primary next-hop of a MHP is same as that of the primary next-hop of the router that was pre-failure optimal point of attachment, an implementation MAY provide the better protection to MHP without incurring any additional computation cost.

2.1.1. IS-IS ATT Bit considerations

Per [RFC1195] a default route needs to be added in Level1 (L1) router to the closest reachable Level1/Level2 (L1/L2) router in the network advertising ATT (attach) bit in its LSP-0 fragment. All L1 routers in the area would do this during the decision process with the next-hop of the default route set to the adjacent router through which the closest L1/L2 router is reachable. The base LFA specification [RFC5286] does not specify any procedure for computing LFA for a default route in IS-IS L1 area. Potentially one MAY consider a default route is being advertised from the boarder L1/L2 router where ATT bit is set and can do LFA computation for the default route. But, when multiple ECMP L1/L2 routers are reachable in an L1 area corresponding best LFAs SHOULD be given for each primary next-hop associated with default route. Considerations as specified in Section 2.1 are applicable for default routes, if the default route is considered as ECMP MHP.

2.2. Links with IGP MAX_METRIC

Section 3.5 and 3.6 of [RFC5286] describes procedures for excluding nodes and links from use in alternate paths based on the maximum link metric (as defined in for IS-IS in [RFC5305] or as defined in [RFC3137] for OSPF). If these procedures are strictly followed, there are situations, as described below, where the only potential alternate available which satisfies the basic loop-free condition will not be considered as alternative.

```
                            +---+  10  +---+  10 +---+
                            | S |------|N1 |-----|D1 |
                            +---+      +---+     +---+
                              |                    |
                           10 |                 10 |
                              |MAX_MET(N2 to S)    |
                              |                    |
                              |       +---+        |
                              +-------|N2 |--------+
                                      +---+
                                   10  |
                                      +---+
                                      |D2 |
                                      +---+
```

Figure 3: Link with IGP MAX_METRIC

In the simple example network, all the link costs have a cost of 10 in both directions, except for the link between S and N2. The S-N2 link has a cost of 10 in the direction from S to N2, and a cost of MAX_METRIC in the direction from N2 to S (0xffffff /2^24 - 1 for IS-IS and 0xffff for OSPF) for a specific end to end Traffic Engineering (TE) requirement of the operator. At node S, D1 is reachable through N1 with cost 20, and D2 is reachable through N2 with cost 20. Even though neighbor N2 satisfies basic loop-free condition (inequality 1 of [RFC5286]) for D1 this could be excluded as potential alternative because of the current exclusions as specified in section 3.5 and 3.6 procedure of [RFC5286]. But, as the primary traffic destined to D2 is continue to use the link and hence irrespective of the reverse metric in this case, the same link MAY be used as a potential LFA for D1.

Alternatively, reverse metric of the link MAY be configured with MAX_METRIC-1, so that the link can be used as an alternative while meeting the TE requirements.

2.3. Multi Topology Considerations

Section 6.2 and 6.3.2 of [RFC5286] state that multi-topology OSPF and ISIS are out of scope for that specification. This memo clarifies and describes the applicability.

In Multi Topology (MT) IGP deployments, for each MT ID, a separate shortest path tree (SPT) is built with topology specific adjacencies, the LFA principles laid out in [RFC5286] are actually applicable for MT IS-IS [RFC5120] LFA SPF. The primary difference in this case is, identifying the eligible-set of neighbors for each LFA computation which is done per MT ID. The eligible-set for each MT ID is determined by the presence of IGP adjacency from Source to the neighboring node on that MT-ID apart from the administrative restrictions and other checks laid out in [RFC5286]. The same is also applicable for OSPF [RFC4915] [MT-OSPF] or different AFs in multi instance OSPFv3 [RFC5838].

However for MT IS-IS, if a default topology is used with MT-ID 0 [RFC5286] and both IPv4 [RFC5305] and IPv6 routes/AFs [RFC5308] are present, then the condition of network congruency is applicable for LFA computation as well. Network congruency here refers to, having same address families provisioned on all the links and all the nodes of the network with MT-ID 0. Here with single decision process both IPv4 and IPv6 next-hops are computed for all the prefixes in the network and similarly with one LFA computation from all eligible neighbors per [RFC5286], all potential alternatives can be computed.

3. IANA Considerations

This document defines no new namespaces and no actions for IANA.

4. Security Considerations

This document does not introduce any new security issues or any change in security considerations as noted in the LFA base specification [RFC5286].

5. Acknowledgements

Authors would like to thank Alia Atlas for detailed review of initial document and providing valuable suggestions. We also thank Bruno Decreane, Stephane Litkowski for their initial review and feedback on the document.

6. References

6.1. Normative References

[RFC1195] Callon, R., "Use of OSI IS-IS for routing in TCP/IP and dual environments", RFC 1195, DOI 10.17487/RFC1195, December 1990, <http://www.rfc-editor.org/info/rfc1195>.

[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>.

6.2. Informative References

[I-D.ietf-rtgwg-lfa-manageability]
Litkowski, S., Decraene, B., Filsfils, C., Raza, K., Horneffer, M., and P. Sarkar, "Operational management of Loop Free Alternates", draft-ietf-rtgwg-lfa-manageability-11 (work in progress), June 2015.

[RFC3137] Retana, A., Nguyen, L., White, R., Zinin, A., and D. McPherson, "OSPF Stub Router Advertisement", RFC 3137, DOI 10.17487/RFC3137, June 2001, <http://www.rfc-editor.org/info/rfc3137>.

[RFC4915] Psenak, P., Mirtorabi, S., Roy, A., Nguyen, L., and P. Pillay-Esnault, "Multi-Topology (MT) Routing in OSPF", RFC 4915, DOI 10.17487/RFC4915, June 2007, <http://www.rfc-editor.org/info/rfc4915>.

[RFC5120] Przygienda, T., Shen, N., and N. Sheth, "M-ISIS: Multi Topology (MT) Routing in Intermediate System to Intermediate Systems (IS-ISs)", RFC 5120, DOI 10.17487/RFC5120, February 2008, <http://www.rfc-editor.org/info/rfc5120>.

[RFC5305] Li, T. and H. Smit, "IS-IS Extensions for Traffic Engineering", RFC 5305, DOI 10.17487/RFC5305, October 2008, <http://www.rfc-editor.org/info/rfc5305>.

[RFC5308] Hopps, C., "Routing IPv6 with IS-IS", RFC 5308, DOI 10.17487/RFC5308, October 2008, <http://www.rfc-editor.org/info/rfc5308>.

[RFC5838] Lindem, A., Ed., Mirtorabi, S., Roy, A., Barnes, M., and R. Aggarwal, "Support of Address Families in OSPFv3", RFC 5838, DOI 10.17487/RFC5838, April 2010, <http://www.rfc-editor.org/info/rfc5838>.

Authors' Addresses

Uma Chunduri
Ericsson Inc.
300 Holger Way,
San Jose, California 95134
USA

Phone: 408 750-5678
Email: uma.chunduri@ericsson.com

Jeff Tantsura
Ericsson Inc.
300 Holger Way,
San Jose, California 95134
USA

Email: jeff.tantsura@ericsson.com

Chris Bowers
Juniper Networks
1194 N. Mathilda Ave.
Sunnyvale, California 94089
USA

Email: cbowers@juniper.net