Network Working Group
Internet-Draft
Expires: September 2, 2010

Clarence Filsfils
Cisco Systems
Pierre Francois
UCLouvain
Mike Shand
Cisco Systems
Bruno Decraene
France Telecom
James Uttaro
ATT
Nicolai Leymann
Martin Horneffer
Deutsche Telekom

March 1, 2010

LFA applicability in SP networks
draft-filsfils-rtgwg-lfa-applicability-00

Abstract

In this draft, we analyze the applicability of LoopFree Alternates in both core and access parts of Service Provider networks. We provide design guides to favor their applicability where relevant, typically in the access part of the network.

Status of this Memo

Table of Contents

1. Introduction

In this document, we analyze the applicability of LoopFree Alternates in both core and access parts of Service Provider networks. We provide design guides to favor their applicability where relevant, typically in the access part of the network.

We first introduce the terminology used in this document in Section 2. In Section 3, we describe typical access network designs and we analyze them for LFA applicability. In Section 4, we describe a simulation framework for the study of LFA applicability in SP core networks, and present results based on various SP networks. We then emphasize the independence between protection schemes used in the core and at the access level of the network. Finally we discuss the key benefits of LFA which stem from its simplicity and we draw some conclusions.

2. Terminology

In this document, we assume that all links to be protected are point-to-point.

We use ISIS as reference. The analysis is equally applicable to OSPF.

A per-prefix LFA for a destination D for a node S is a precomputed backup IGP nexthop for that destination. This backup IGP nexthop can be link protecting or node protecting.

Link-protecting: A neighbor N is a link-protecting per-prefix LFA for S's route to D if equation eq1 is satisfied, with eq1 == ND < NS + SD where XY refers to the IGP distance from X to Y. This is in line with the definition of an LFA in [RFC5714].

Node-protecting: A Neighbor N is a node-protecting LFA for S's route to D, with initial IGP nexthop F if N is a link-protecting LFA for D and equation eq2 is satisfied, with eq2 == ND < NF + FD. This is in line with the definition of a Node-Protecting Alternate Next-Hop in [RFC5714].

De facto node-protecting LFA: this is a link-protecting LFA that turns out to be node-protecting. This occurs in cases illustrated by the following examples :

- o The LFA candidate that is picked by S actually satisfies Equation eq2 but S did not verify that property. The show command issued by the operator would not indicate this LFA as "node protecting" while in practice (de facto) it is.

o A cascading effect of multiple LFA's can also provide de facto node protection. Equation eq2 is not satisfied, but the combined activation of LFAs by some other neighbors of the failing node F provides (de facto) node protection. In other words, it puts the dataplane in a state such that packets forwarded by S ultimately reach a neighbor of F that has a node-protecting LFA. Note that in this case S cannot indicate the node-protecting behavior of the repair without running additional computations.

Per-Link LFA: a per-link LFA for the link SF is one precomputed backup IGP nexthop for all the destinations reached through SF. This is a neighbor of the repairing node that is a per-Prefix LFA for all the prefixes that the repairing node reaches through SF. Note that such a per-link LFA exists if S has a per-prefix LFA for destination F.

```
      D
     / \
 10 /   \ 10
   /     \
  G       H----------.
  |       |          |
1 |     1 |          |
  |       |          |
  B       C          | 10
  |       |\         |
  |       | \        |
  |       |  \ 6     |
  |       |   \      |
7 |    10 |    E      F
  |       |   /     /
  |       |  / 6   / 5
  |       | /     /
  |       |/     /
  A-------S-----/
      7
```

Figure 1: Example 1

In Figure 1, considering the protection of link SC, we can see that A, E, and F are per-prefix LFAs for destination D, as none of them use S to reach D.

For destination D, A and F are node-protecting LFA as they do not reach D through node C, while E is not node-protecting for S as it reaches D through C.

If S does not compute and select node-protecting LFAs, there is a chance that S picks the non node-protecting LFA E, although A and F were node-protecting LFAs. If S enforces the selection of node-protecting LFAs, then in the case of the single failure of link SC, S will first activate its LFA and deviate traffic addressed to D along S-A-B-G-D and/or S-F-H-D, and then converge to its post-convergence optimal path S-E-C-H-D.

A is not a per-link LFA for link SC because A reaches C via S. E is a per-Link LFA for link SC as it reaches C through link EC. This per-link LFA does not provide de facto node protection. Upon failure of node C, S would fast-reroute D-destined packets to its per-link lfa (= E). E would himself detect the failure of EC and hence activate its own per-link LFA (=S). Traffic addressed to D would be trapped in a loop and hence there is no de facto node protection behavior.

If there were a link between E and F, that E would pick as its LFA for destination D, then E would provide de facto node protection for S, as upon the activation of its LFA, S would deviate traffic addressed to D towards E, which in turns deviates that traffic to F, which does not reach D through C.

F is a per-Link LFA for link SC as F reaches C via H. This per-link LFA is de facto node-protecting for destination D as F reaches D via F-H-D.

MicroLoop (uloop): the occurrence of a transient forwarding loop during a routing transition (as defined in [RFC5714]).

In Figure 1, the loss of link SE cannot create any uloop because: 1/The link is only used to reach destination E and 2/ S is the sole node changing its path to E upon link SE failure. 3/ S's shortest path to E after the failure goes via C. 4/C's best path to E (before and after link SC failure) is via CE.

To the contrary, upon failure of link AB, a microloop may form for traffic destined to B. Indeed, if A updates its FIB before S, A will deviate B-destined traffic towards S, while S is still forwarding this traffic to A.

3. Access Network

The access part of the network often represents the majority of the nodes and links. It is organized in several tens or more of regions interconnected by the core network. Very often the core acts as an ISIS level2 domain (OSPF area 0) while each access region is confined in an ISIS level1 domain (OSPF non0 area). Very often, the network

topology within each access region is derived from a unique template common across the whole access network. Within an access region itself, the network is made of several aggregation regions, each following the same interconnection topologies.

For these reasons, we base the analysis of the LFA applicability in the access network on the following abstract model:

- o We analyze a single access region.
- o Two routers (C1 and C2) provide connectivity between the access region and the rest of the network. If a link connects these two routers in the region area, then it has a symmetric IGP metric c.
- o We analyze a single aggregation region within the access region. Two aggregation routers (A1 and A2) interconnect the aggregation region to the two routers C1 and C2 for the analyzed access region. If a link connects A1 to A2 then it has a symmetric IGP metric a. If a link connects an A to a C router then, for sake of generality, we will call d the metric for the directed link CA and u the metric for the AC directed link.
- o We analyze two edge routers E1 and E2 in the access region. Each is either dual-homed directly into C1 and C2 xor into A1 and A2. The directed link metric between Cx/Ax and Ey is d and u in the opposite direction.
- o We assume a multi-level IGP domain. The analyzed access region forms a level-1 domain. The core is the level-2 domain. We assume that the link C is L1L2. We assume that the loopbacks of the C routers are part of the L2 topology. L1 routers learn about them as propagated routes (L2=>L1 with Down bit set). We remind that if an L1L2 router learns about X/x as an L1 path P1, an L2 path P2 and an L1L2 path P12, then it will prefer path P1. If P1 is lost, then it will prefer path P2.
- o We assume that all the C, A and E routers may be connected to customers and hence we analyze LFA coverage for the loopbacks of each type of node.
- o We assume that no useful traffic is directed to router-to-router subnets and hence we do not analyze LFA applicability for these.
- o A prefix P models an important IGP destination that is not present in the local access region. The igp metric from C1 to P is x and the metric from C2 to P is x+e.
- o We analyze LFA coverage against all link and node failures within the access region.
- o WxYz refers to the link from Wx to Yz.
- o We assume that c < d + u and a < d + u (commonly agreed design rule).
- o In the square access design, we assume that c < a (commonly agreed design rule).
- o We analyze the most frequent topologies found in an access region.

- o We first analyze per-prefix LFA applicability and then per-link.
- o The topologies are symmetric with respect to a vertical axe and hence we only detail the logic for the link and node failures of the left half of the topology.
- o We do not consider SRLGs. Future revisions of the draft will address this topic.

3.1. Triangle

We describe the LFA applicability for the failures of each direction of link C1E1, E1 and C1 (Figure 2), and for the failure of each node.

```
          P
         / \
       x/   \x+e
       /     \
     C1--c--C2
       |\    /|
    d/u|  \/  |d/u
       | /  \ |
       E1     E2
```

Figure 2: Triangle

3.1.1. E1C1 failure

3.1.1.1. Per-Prefix LFA

Three destinations are impacted by this link failure: C1, E2 and P.

The LFA for destination C1 is C2 because eq1 == c < d + u. Node protection for route C1 is not applicable. (if C1 goes down, traffic destined to C1 is lost anyway).

The LFA to E2 is via C2 because eq1 == d < d+u+d. It is node protecting because eq2 == d < c + d.

The LFA to P is via C2 because eq1 == c < d + u. It is node protecting if eq2 == x + e < x + c, i.e., if e < c. This relationship between e and c is an important aspect of the analysis, which is discussed in detail in Section 3.5 and Section 3.6

Conclusion: all important intra-PoP routes with primary interface E1C1 benefit from LFA link and node protection. All important inter-PoP routes with primary interface E1C1 benefit from LFA link protection, and also from node protection if e < c.

3.1.1.2. Per-Link LFA

We have a per-prefix LFA to C1 and hence we have a per-link LFA for link E1C1. All impacted destinations are protected for link failure. In case of C1 node failure, the traffic to C1 is lost (by definition), the traffic to E2 is de facto protected against node failure and the traffic to P is de facto protected when e < c.

3.1.2. C1E1 failure

3.1.2.1. Per-Prefix LFA

C1 has one single primary route via C1E1: the route to E1 (because c < d + u).

C1's LFA to E1 is via C2 because eq1 == d < c + d.

Node protection upon E1's failure is not applicable as the only impacted traffic is sinked at E1 and hence is lost anyway.

Conclusion: all important routes with primary interface C1E1 benefit from LFA link protection. Node protection is not applicable.

3.1.2.2. Per-Link LFA

We have a per-prefix LFA to E1 and hence we have a per-link LFA for link C1E1. De facto node protection is not applicable.

3.1.3. uLoop

The IGP convergence cannot create any uloop. See Section 3.7.

3.1.4. Conclusion

All important intra-PoP routes benefit from LFA link and node protection or de facto node protection. All important inter-PoP routes benefit from LFA link protection. De facto node protection is ensured if e < c (this is particularly the case for dual-plane core or two-tiered-igp-metric design, see later sections).

The IGP convergence does not cause any uLoop.

Per-link LFA and per-Prefix LFA provide the same protection benefits.

3.2. Full-Mesh

We describe the LFA applicability for the failures of C1A1, A1E1, E1, A1 and C1 (Figure 3).

```
                                 P
                                / \
                             x/   \x+e
                             /      \
                           C1--c--C2
                             |\   /|
                             | \ / |
                       d/u |  \  | d/u
                             | / \ |
                             |/   \|
                           A1--a--A2
                             |\    /|
                         d/u|  \/  |d/u
                             | /  \ |
                            E1      E2
```

Figure 3: Full-Mesh

3.2.1. E1A1 failure

3.2.1.1. Per-Prefix LFA

Four destinations are impacted by this link failure: A1, C1, E2 and P.

The LFA for A1 is A2: eq1 == a < d + u. Node protection for route A1 is not applicable (if A1 goes down, traffic to A1 is lost anyway).

The LFA for C1 is A2: eq1 == u < d + u + u. Node protection for route C1 is guaranteed: eq2 == u < a + u.

The LFA to E2 is via A2: eq1 == d < d+u+d. Node protection is guaranteed: eq2 == d < a + d.

The LFA to P is via A2: eq1 == u + x < d + u + u + x. Node protection is guaranteed: eq2 == u+ x < a + u + x.

Conclusion: all important intra-PoP and inter-PoP routes with primary interface E1A1 benefit from LFA link and node protection.

3.2.1.2. Per-Link LFA

We have a per-prefix LFA to A1 and hence we have a per-link LFA for link E1A1. All impacted destinations are protected for link failure. De facto node protection is provided for all prefixes (except to A1 which is not applicable).

3.2.2. A1E1 failure

3.2.2.1. Per-Prefix LFA

A1 has one single primary route via A1E1: the route to E1 (because c < d + u).

A1's LFA to E1 is via A2: eq1 == d < a + d.

Node protection upon E1's failure is not applicable as the only impacted traffic is sinked at E1 and hence is lost anyway.

Conclusion: all important routes with primary interface A1E1 benefit from LFA link protection. Node protection is not applicable.

3.2.2.2. Per-Link LFA

We have a per-prefix LFA to E1 and hence we have a per-link LFA for link C1E1. De facto node protection is not applicable.

3.2.3. A1C1 failure

3.2.3.1. Per-Prefix LFA

Two destinations are impacted by this link failure: C1 and P.

The LFA for C1 is C2 because eq1 == c < d + u. Node protection for route C1 is not applicable (if C1 goes down, traffic to C1 is lost anyway).

The LFA for P is via C2 because eq1 == c < d + u. It is de facto protected for node failure if eq2 == x + e < x + c.

Conclusion: all important intra-PoP routes with primary interface A1C1 benefit from LFA link protection (node protection is not applicable). All important inter-PoP routes with primary interface E1C1 benefit from LFA link protection (and from de facto node protection if e < c).

3.2.3.2. Per-Link LFA

We have a per-prefix LFA to C1 and hence we have a per-link LFA for link A1C1. All impacted destinations are protected for link failure. In case of C1 node failure, the traffic to C1 is lost (by definition) and the traffic to P is de facto node protected if e < c.

3.2.4. C1A1 failure

3.2.4.1. Per-Prefix LFA

C1 has three routes via C1A1: A1, E1 and E2. E2 behaves like E1 and hence is not analyzed further.

C1's LFA to A1 is via C2 because we assumed c < a and eq1 == d < c + d. Node protection upon A1's failure is not applicable as the traffic to A1 is lost anyway.

C1's LFA to E1 is via A2: eq1 == d < u+ d + d. Node protection upon A1's failure is guaranteed because: eq2 == d < a + d.

Conclusion: all important routes with primary interface C1A1 benefit from LFA link protection. Node protection is guaranteed where applicable.

3.2.4.2. Per-Link LFA

We have a per-prefix LFA to A1 and hence we have a per-link LFA for link C1E1. De facto node protection is available.

3.2.5. uLoop

The IGP convergence cannot create any uloop. See Section 3.7.

3.2.6. Conclusion

All important intra-PoP routes benefit from LFA link and node protection.

All important inter-PoP routes benefit from LFA link protection. They benefit from node protection upon failure of A nodes. They benefit from node protections upon failure of C nodes if e < c (this is particularly the case for dual-plane core or two-tiered-igp-metric design, see later sections).

The IGP convergence does not cause any uLoop.

Per-link LFA and per-Prefix LFA provide the same protection benefits.

3.3. Square

We describe the LFA applicability for the failures of C1A1, A1E1, E1, A1 and C1 (Figure 4).

```
          P
         / \
       x/   \x+e
       /     \
      C1--c--C2
       |\    | \
       | \   |  +-------+
   d/u |  \  |           \
       |   +-|-----+      \
       |     |      \       \
      A1--a--A2     A3--a--A4
       |\    /|      |    /
    d/u|  \/  |d/u   |  /
       | /  \ |      |/
       E1     E2     E3
```

Figure 4: Square

3.3.1. E1A1 failure

3.3.1.1. Per-Prefix LFA

E1 has six routes via E1A1: A1, C1, P, E2, A3, E3.

E1's LFA route to A1 is via A2 because eq1 == a < d + u. Node protection for traffic to A1 upon A1 node failure is not applicable.

E1's LFA route to A3 is via A2 because eq1 == u + c + d < d + u + u + d. This LFA is guaranteed to be node protecting because eq2 == u + c + d < a + u + d.

E1's LFA route to C1 is via A2 because eq1 == u + c < d + u + u. This LFA is guaranteed to be node protecting because eq2 == u + c < a + u.

E1's primary route to E2 is via ECMP(E1A1, E1A2). The LFA for the first ECMP path (via A1) is the second ECMP path (via A2). This LFA is guaranteed to be node protecting because eq2 == d < a + d.

E1's primary route to E3 is via ECMP(E1A1, E1A2). The LFA for the first ECMP path (via A1) is the second ECMP path (via A2). This LFA is guaranteed to be node protecting because eq2 == u + d + d < a + u d + d.

If e=0: E1's primary route to P is via ECMP(E1A1, E1A2). The LFA for the first ECMP path (via A1) is the second ECMP path (via A2). This LFA is guaranteed to be node protecting because eq2 == u + x + 0 < a + u + x .

If e<>0: E1's primary route to P is via E1A1. Its LFA is via A2 because eq1 == u + c + x < d + u + u + x. This LFA is guaranteed to be node protecting because eq2 == u + c + x < a + u + x.

Conclusion: all important intra-PoP and inter-PoP routes with primary interface E1A1 benefit from LFA link protection and node protection.

3.3.1.2. Per-Link LFA

We have a per-prefix LFA for A1 and hence we have a per-link LFA for link E1A1. All important intra-PoP and inter-PoP routes with primary interface E1A1 benefit from LFA per-link protection and de facto node protection.

3.3.2. A1E1 failure

3.3.2.1. Per-Prefix LFA

A1 has one single primary route via A1E1: the route to E1.

A1's LFA for route E1 is the path via A2 because eq1 == d < a + d. Node protection is not applicable.

Conclusion: all important routes with primary interface A1E1 benefit from LFA link protection. Node protection is not applicable.

3.3.2.2. Per-Link LFA

All important routes with primary interface A1E1 benefit from LFA link protection. De facto node protection is not applicable.

3.3.3. A1C1 failure

3.3.3.1. Per-Prefix LFA

Four destinations are impacted when A1C1 fails: C1, A3, E3, and P.

A1's LFA to C1 is via A2 because eq1 == u + c < a + u. Node protection property is not applicable for traffic to C1 when C1 fails.

A1's LFA to A3 is via A2 because eq1 == u + c + d < a + u + d. It is de facto node protecting as a < u + c + d (as we assumed a < u + d). Indeed A2 forwards traffic destined to A3 to C2, and C2 has a node protecting LFA for A3 w.r.t the failure of C2C1, being A4, as a < u + c + d. Hence the cascading application of LFAs by A1 and C2 during the failure of C1 provides de facto node protection.

A1's LFA to E3 is via A2 because eq1 == u + d + d < a + u + d + d. It is node protecting because eq2 == u + d + d < u + c + d + d.

A1's primary route to P is via C1 (even if e=0, u+x < u + c + x). The LFA is via A2 because eq1 == [u + c + x < a + u + x]. This LFA is node protecting (from the viewpoint of A1 computing eq2) if eq2 == u + x + e < u + c + x hence if e < c.

Conclusion: all important intra-PoP routes with primary interface A1C1 benefit from LFA link protection and node protection. Note that A3 benefits from a de facto node protection. All important inter-PoP routes with primary interface A1C1 benefit from LFA link protection. They also benefit from node protection if e < c.

3.3.3.2. Per-Link LFA

All important intra-PoP routes with primary interface A1C1 benefit from LFA link protection and de facto node protection. All important inter-PoP routes with primary interface A1C1 benefit from LFA link protection. They also benefit from de facto node protection if e < c.

3.3.4. C1A1 failure

3.3.4.1. Per-Prefix LFA

Three destinations are impacted by C1A1 link failure: A1, E1 and E2. E2's analysis is the same as E1 and hence is omitted.

C1's has no LFA for A1. Indeed, all its neighbors (C2 and A3) have a shortest path to A1 via C1. This is due to the assumption (c < a).

C1's LFA for E1 is via C2 because eq1 == d + d < c + d + d. It provides node protection because eq2 == d + d < d + a + d.

Conclusion: all important intra-PoP routes with primary interface A1C1 except A1 benefit from LFA link protection and node protection.

3.3.4.2. Per-Link LFA

C1 does not have a per-prefix LFA for destination A1 and hence there is no per-link LFA for the link C1A1.

3.3.4.3. Assumptions on the values of c and a

If c > a, then C1 would have a per-prefix LFA for A1 and hence link C1A1 would have a per-link LFA. However, in that case, A1 would no longer have a per-prefix LFA for C1 and hence A1 would no longer have

a per-link LFA for the link A1C1.

The commonly agreed design rule (c < a) is beneficial for a deployment using per-link LFA: it provides a per-link LFA for the most important direction (A1C1). Indeed, there are many more prefixes reachable over A1C1 then over C1A1. As the IGP convergence duration is proportional to the number of routes to update, there is a better benefit in leveraging LFA FRR for the link A1C1 than the link C1A1.

Note as well that the consequence of this assumption is much more important for per-link LFA than for per-prefix LFA.

For per-prefix LFA, in case of link C1A1 failure, we do have a per-prefix LFA for E1, E2 and any node subtended below A1 and A2. Typically most of the traffic traversing the link C1A1 is directed to these E nodes and hence the lack of per-prefix LFA for the destination A1 might be insignificant. This is a good example of the coverage benefit of per-prefix LFA over per-link LFA.

Finally note that c = a is the worst choice as in this case there C1 has no per-prefix LFA for A1 (and vice versa) and hence there is no per-link LFA for C1A1 and A1C1.

3.3.5. Conclusion

All important intra-PoP routes benefit from LFA link and node protection with one exception: C1 has no per-prefix LFA to A1.

All important inter-PoP routes benefit from LFA link protection. They benefit from node protection if e < c.

Per-link LFA provides the same protection coverage as per-prefix LFA with two exceptions. First, C1A1 has no per-link LFA at all. Second, when per-prefix LFA provides node protection (eq2 is satisfied), per-link LFA provides effective de facto node protection.

3.3.6. A square might become a full-mesh

If the vertical links of the square are made of parallel links (at L3 or at L2), then one should consider splitting these "vertical links" into "vertical and crossed links". The topology becomes "full-mesh". One should also ensure that the two resulting set of links (vertical and crossed) do not share any SRLG.

A typical reason preventing this is that the A1C1 bandwidth may be within a building while the A1C2 is between buildings. Hence while from a router port viewpoint the operation is cost-neutral, it is not

from a cost of bandwidth viewpoint.

3.3.7. A full-mesh might be more economical than a square

In a full-mesh, the vertical and cross-links play the dominant role as they support most of the primary and backup paths. The capacity of the horizontal links can be dimensioned on the basis of traffic destined to a single C or a single A and a single E node.

3.4. Extended U

For the Extended U topology, we define the following terminology:

C1L1: the node "C1" as seen in topology L1.

C1L2: the node "C1" as seen in topology L2.

C1LO: the loopback of C1. This loopback is in L2.

Let us also remind that C1 and C2 are L1L2 routers and that their loopbacks are in L2 only.

```
                                     P
                                    / \
                                  x/   \x+e
                                  /     \
                                 C1<...>C2
                                  |\    | \
                                  | \   |  +-------+
                              d/u |  \  |           \
                                  |   +-|-----+      \
                                  |     |      \      \
                                 A1--a--A2     A3--a--A4
                                  |\    /|      |    /
                               d/u|  \/  |d/u   |  /
                                  | /  \ |      |/
                                  E1     E2     E3
```

Figure 5: Extended U

There is no L1 link between C1 and C2. There might be an L2 link between C1 and C2. This is not relevant as this is not seen from the viewpoint of the L1 topology which is the focus of our analysis.

It is guaranteed that there is a path from C1L0 to C2LO within the L2 topology (except if the L2 topology partitions which is very unlikely and hence not analyzed here). We call "c" its path cost. Once again, we assume that c < a.

We exploit this property to create a tunnel T between C1LO and C2LO. Once again, as the source and destination addresses are the loopbacks of C1 and C2 and these loopbacks are in L2 only, it is guaranteed that the tunnel does not transit via the L1 domain.

ISIS does not run over the tunnel and hence the tunnel is not used for any primary paths within the L1 or L2 topology.

Within topology Level1, we configure C1 (C2) with a Level1 LFA extended neighbor "C2 via tunnel T" ("C1 via tunnel T").

A router supporting such extension learns that it has one additional potential neighbor in topology Level1 when checking for LFA's.

The L1 topology learns about C1LO as an L2=>L1 route with Down bit set propagated by C1L1 and C2L1. The metric advertised by C2L1 is bigger than the metric advertised by C1L1 by "c".

The L1 topology learns about P as an L2=>L1 routes with Down bit set propagated by C1L1 and C2L1. The metric advertised by C2L1 is bigger than the metric advertised by C1L1 by "e". This implies that e <= c.

3.4.1. E1A1 failure

3.4.1.1. Per-Prefix LFA

Five destinations are impacted by E1A1 link failure: A1, C1LO, E2, E3 and P.

The LFA for A1 is via A2 because eq1 == a < d + u. Node protection for traffic to A1 upon A1 node failure is not applicable.

The LFA for E2 is via A2 because eq1 == d < d + u + d. Node protection is guaranteed because eq2 == d < a + d.

The LFA for E3 is via A2 because eq1 == u + d + d < d + u + d + d. Node protection is guaranteed because eq2 == u + d + d < a + u + d + d.

The LFA for C1LO is via A2 because eq1 == u + c < d + u + u. Node protection is guaranteed because eq2 == u + c < a + u.

If e=0: E1's primary route to P is via ECMP(E1A1, E1A2). The LFA for the first ECMP path (via A1) is the second ECMP path (via A2). Node protection is possible because eq2 == u + x < a + u + x.

If e<>0: E1's primary route to P is via E1A1. Its LFA is via A2 because eq1 == a + c + x < d + u + u + x. Node protection is

guaranteed because eq2 == u + x + e < a + u + x <=> e < a. This is true because e <= c and c < a.

Conclusion: same as the square topology.

3.4.1.2. Per-Link LFA

Same as the square topology.

3.4.2. A1E1 failure

3.4.2.1. Per-Prefix LFA

Same as the square topology.

3.4.2.2. Per-Link LFA

Same as the square topology.

3.4.3. A1C1 failure

3.4.3.1. Per-Prefix LFA

Three destinations are impacted when A1C1 fails: C1, E3 and P.

A1's LFA to C1LO is via A2 because eq1 == u + c < a + u. Node protection property is not applicable for traffic to C1 when C1 fails.

A1's LFA to E3 is via A2 because eq1 == u + d + d < d + u + u + d + d. Node protection is guaranteed because eq2 == u + d + d < a + u + d + d.

A1's primary route to P is via C1 (even if e=0, u + x < a + u + x). The LFA is via A2 because eq1 == u + x + e < a + u + x <=> e < a (which is true see above). Node protection is guaranteed because eq2 == u + x + e < a + u + x.

Conclusion: same as the square topology

3.4.3.2. Per-Link LFA

Same as the square topology.

3.4.4. C1A1 failure

3.4.4.1. Per-Prefix LFA

Three destinations are impacted by C1A1 link failure: A1, E1 and E2. E2's analysis is the same as E1 and hence is omitted.

C1L1 has an LFA for A1 via the extended neighbor C2L1 reachable via tunnel T. Indeed, eq1 is true: d + a < d + a + u + d. From the viewpoint of C1L1, C2L1's path to C1L1 is C2L1-A2-A1-C1L1. Remember the tunnel is not seen by ISIS for computing primary paths! Node protection is not applicable for traffic to A1 when A1 fails.

C1L1's LFA for E1 is via extended neighbor C2L1 (over tunnel T) because eq1 == d + d < d + a + u + d + d. Node protection is guaranteed because eq2 == d + d < d + a + d.

3.4.4.2. Per-Link LFA

C1 has a per-prefix LFA for destination A1 and hence there is a per-link LFA for the link C1A1. Node resistance is applicable for traffic to E1 (and E2).

3.4.5. Conclusion

The extended U topology is as good as the square topology.

It does not require any cross links between the A and C nodes within an aggregation region. It does not need an L1 link between the C routers in an access region. Note that a link between the C routers might exist in the L2 topology.

3.5. Dual-plane Core and its impact on the Access LFA analysis

A Dual-plane core is defined as follows

- o Each access region k is connected to the core by two C routers (C(1,k) and C(2,k)).
- o C(1,k) is part of Plane1 of the dual-plane core.
- o C(2,k) is part of Plane2 of the dual-plane core.
- o C(1,k) has a link to C(2, l) iff k = l
- o {C(1,k) has a link to C(1, l)} iff {C(2,k) has a link to C(2, l)}

In a dual-plane core design, e = 0 and hence the LFA node-protection coverage is improved in all the analyzed topologies.

3.6. Two-tiered IGP metric allocation

A Two-tiered IGP metric allocation scheme is defined as follows

- o all the link metrics used in the L2 domain are part of range R1
- o all the link metrics used in an L1 domain are part of range R2
- o range R1 << range R2 such that the difference e = C2P - C1P is smaller than any link metric within an access region.

Assuming such an IGP metric allocation, the following properties are guaranteed: c < a and e < c.

3.7. uLoop analysis

In this section, we analyse a case where the routing transition following the failure of a link may have some uLoop potential for one destination. Then we show that all the other cases do not have uLoop potential.

In the square design, upon the failure of link C1A1, traffic addressed to A1 can undergo a transient forwarding loop as C1 reroutes traffic to C2, which initially reaches A1 through C1, as c < a. This loop will actually occur when C1 updates its FIB for destination A1 before C2.

It can be shown that all the other routing transitions following a link failure in the analyzed topologies do not have uLoop potential. Indeed, in each case, for all destinations affected by the failure, the rerouting nodes deviate their traffic directly to adjacent nodes whose paths towards these destinations do not change. As a consequence, all these routing transitions cannot undergo transient forwarding loops.

For example, in the square topology, the failure of directed link A1C1 does not lead to any uloop. The destinations reached over that directed link are C1 and P. A1 and E1's shortest paths to these destinations after the convergence go via A2. A2's path to C1 and P is not using A1C1 before the failure, hence no uloop may occur.

3.8. Summary

1. **Intra Area Destinations**

    Link Protection
    - + Triangle: Full
    - + Full-Mesh: Full
    - + Square: Full, except C1 has no LFA for dest A1
    - + Extended U: Full

    Node Protection
    - + Triangle: Full
    - + Full-Mesh: Full
    - + Square: Full

         + Extended U: Full
   2. Inter Area Destinations
         Link Protection
         + Triangle: Full
         + Full-Mesh: Full
         + Square: Full
         + Extended U: Full
         Node Protection
         + Triangle: yes if e<c
         + Full-Mesh: yes for A failure, if e<c for C failure
         + Square: yes for A failure, if e<c for C failure
         + Extended U : yes if e<= c and c < a
   3. ULoops
       * Triangle: None
       * Full-Mesh: None
       * Square: None, except traffic to A1 when C1A1 fails
       * Extended U : None, if a > e
   4. Per-Link LFA vs Per-Prefix LFA
       * Triangle: Same
       * Full-Mesh: Same
       * Square: Same except C1A1 has no per-Link LFA. In practice, this means that per-prefix LFAs will be used (hence C1 has no LFA for dest=E1 and dest=A1)
       * Extended U : Same

4. Core Network

In the backbone, the optimization of the network design to achieve the maximum LFA protection is less straightforward than in the case of the access/aggregation network.

The main optimization objectives for backbone topology design are cost, latency, and bandwidth, constrained by the availability of fiber. Optimizing the design for Local IP restoration is more likely to be considered as a non-primary objective. For example, the way the fiber is laid out and the resulting cost to change it leads to ring topologies in some backbone networks.

Also, the capacity planning process is already complex in the backbone. It needs to make sure that the traffic matrix (demand) is supported by the underlying network (capacity) under all possible variation of the underlying network (what-if scenario related to one-srlg failure). Classically, "supported" means that no congestion be experienced and that the demands be routed along the appropriate latency paths. Selecting LFA as a deterministic FRR solution for the backbone would require to enhance the capacity planning process to add a third constraint: each variation of the underlying network

should lead to a sufficient LFA coverage.

To the contrary, the access network is based on many replications of a small number of well-known (well-engineered) topologies. The LFA coverage is deterministic and is independent of additions/insertions of a new edge device, a new aggregation sub-region or a new access region.

In practice, we believe that there are three profiles for the backbone applicability of LFA.

In the first profile, the designer plans all the network resilience on IGP convergence. In such case, LFA is a free bonus. If an LFA is available, then the loss of connectivity is likely reduced by a factor 10 (50msec vs 500msec), else the loss of connectivity depends on IGP convergence which is anyway the initial target. LFA should be very successful here as it provides a significant improvement without any additional cost.

In the second profile, the designer seeks a very high and deterministic FRR coverage and he either does not want or cannot engineer the topology. LFA should not be considered in this case. MPLS TE FRR would perform much better in this environment. Explicit routing ensures that a backup path exists what-ever the underlying topology.

In the third profile, the designer seeks a very high and deterministic FRR coverage and he does engineer the topology. LFA is appealing in this scenario as it can provide a very simple way to obtain protection.

For the reasons explained previously, the backbone applicability should be analyzed on a case by case basis and it is difficult to derive generic rules.

In order to help the reader to assess the LFA applicability in its own case, we provide in the next section some simulation results based on 11 real backbone topologies.

4.1. Simulation Framework

We usually receive the complete ISIS/OSPF linkstate database taken on a core router. We parse it to obtain the topology. During this process, we eliminate all nodes connected to the topology with a single link and all prefixes except a single "node address" per router. We compute the availability of per-prefix LFA's to all these node addresses which we call "destinations" hereafter. We treat each link in each direction.

For each (directed) link, we compute whether we have a per-prefix LFA to the next-hop. If so, we have a per-link LFA for the link.

The Per-link-LFA coverage for a topology T is the ratio of the number of links with a per-link LFA divided by the total number of links.

For each link, we compute the number of destinations whose primary path involves the analyzed link. For each such destination, we compute whether a per-prefix LFA exists.

The Per-Prefix-LFA coverage for a topology T is the ratio:

(the sum across all links of the number of destinations with a primary path over the link and a per-prefix LFA)

divided by

(the sum across all links of the number of destinations with a primary path over the link)

4.2. Data Set

Our data set is based on 11 SP core topologies with different geographical scopes: worldwide, national and regional. The number of nodes range from 600 to 16. The average link-to-node ratio is 2.3 with a minimum of 1.2 and maximum of 6.

4.3. Simulation results

Topology	Per-link LFA	Per-prefix LFA
T1	45%	77%
T2	49%	99%
T3	88%	99%
T4	68%	84%
T5	75%	94%
T6	87%	99%
T7	16%	67%
T8	87%	100%
T9	67%	80%
T10	98%	100%
T11	59%	77%
Average	67%	89%
Median	68%	94%

Table 1: Core LFA Coverages

In Table 1, we observe a wide variation in terms of LFA coverage across topologies; From 67% to 100% for the per-prefix LFA coverage, and from 16% to 98% for the per-link LFA coverage. Several topologies have been optimized for LFAs (T3, 6, 8 and 10). This illustrates the need for case by case analysis when considering LFA for core networks.

It should be noted that, to the contrary of the access/aggregation topologies, per-prefix LFA outperforms per-link LFA in the backbone.

5. Core and Access protection schemes are independent

Specifically, a design might use LFA FRR in the access and MPLS TE FRR in the core.

LFA provides great benefits for the access network due to its excellent access coverage and its simplicity.

MPLS TE FRR's topology independence might prove beneficial in the core when either the LFA FRR coverage is judged too small and/or the designer feels unable to optimize the topology to improve the LFA coverage.

6. Simplicity and other LFA benefits

The LFA solution provides significant benefits which mainly stem from its simplicity.

The LFA behavior is an automated process that makes fast restoration an intrinsic part of the IGP, with no additional configuration burden in the IGP or any other protocol.

Thanks to this integration, the use of multiple areas in the IGP does not make Fast Restoration more complex to achieve than in a single area IGP design.

There is no requirement for network-wide upgrade as LFAs do not require any protocol change and hence can be deployed router by router.

With LFAs, the backup paths are pre-computed and installed in the dataplane in advance of the failure. Assuming a fast enough FIB update time compared to the total number of (important) prefixes, a "<50msec repair" requirement becomes achievable. With a prefix-independent implementation, LFAs have a fixed repair time, as it only depends on the failure detection time and the time to activate the

LFA behavior, which does not scale with the number of prefixes to be fast rerouted.

Link and node protection are provided together and without operational difference (as a comparison, MPLS TE FRR link and node protections require different types of backup tunnels and different grades of operational complexity).

The per-prefix mode of LFAs allows for a simpler and more efficient capacity planning. As the backup path of each prefix is optimized individually, the load to be fast rerouted can be spread on a set of shortest-repair-paths (as opposed to one single backup tunnel). This leads for a simpler and more efficient capacity planning process that takes congestion during protection into account.

7. Security Considerations

This document does not introduce any new security considerations.

8. IANA considerations

This draft does not require any IANA considerations.

9. Conclusions

LFA is an important protection alternative for IP/MPLS networks.

Its simplicity benefit is significant, in terms of automation and integration with the default IGP behavior and the abscense of any requirement for network-wide upgrade. The technology does not require any protocol change and hence can be deployed router by router.

At first sight, these significant simplicity benefits are negated by the topological dependency of its applicability.

The purpose of this document was to highlight that very frequent access and aggregation topologies benefit from excellent link and node LFA coverage.

A second objective consisted in describing the three different profiles of LFA applicability for the IP/MPLS core networks and illustrating them with simulation results based on real SP core topologies.

Future versions of this document will cover additional access topologies and will describe multicast applicability.

10. References

[RFC5714] Shand, M. and S. Bryant, "IP Fast Reroute Framework", RFC 5714, January 2010.

Authors' Addresses

Clarence Filsfils
Cisco Systems
Brussels 1000
BE

Email: cf@cisco.com

Pierre Francois
UCLouvain
Place Ste Barbe, 2
Louvain-la-Neuve 1348
BE

Email: pierre.francois@uclouvain.be
URI: http://inl.info.ucl.ac.be/pfr

Mike Shand
Cisco Systems
Green Park, 250, Longwater Avenue,
Reading RG2 6GB
UK

Email: mshand@cisco.com

Bruno Decraene
France Telecom
38-40 rue du General Leclerc
92794 Issi Moulineaux cedex 9
FR

Email: bruno.decraene@orange-ftgroup.com

James Uttaro
ATT
200 S. Laurel Avenue
Middletown, NJ 07748
US

Email: uttaro@att.com

Nicolai Leymann
Deutsche Telekom
Winterfeldtstrasse 21
Berlin 10781
DE

Email: nicolai.leymann@t-systems.com

Martin Horneffer
Deutsche Telekom
Hammer Str. 216-226
Muenster 48153
DE

Email: Martin.Horneffer@t-com.net