

Routing Area Working Group  
Internet-Draft  
Intended status: Standards Track  
Expires: July 17, 2015

S. Litkowski  
B. Decraene  
Orange  
C. Filsfils  
K. Raza  
Cisco Systems  
M. Horneffer  
Deutsche Telekom  
P. Sarkar  
Juniper Networks  
January 13, 2015

Operational management of Loop Free Alternates  
draft-ietf-rtgwg-lfa-manageability-07

## Abstract

Loop Free Alternates (LFA), as defined in [RFC 5286](#) is an IP Fast ReRoute (IP FRR) mechanism enabling traffic protection for IP traffic (and MPLS LDP traffic by extension). Following first deployment experiences, this document provides operational feedback on LFA, highlights some limitations, and proposes a set of refinements to address those limitations. It also proposes required management specifications.

This proposal is also applicable to remote LFA solution.

## 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](#)].

## 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."

Internet-Draft

LFA manageability

January 2015

This Internet-Draft will expire on July 17, 2015.

## Copyright Notice

Copyright (c) 2015 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to [BCP 78](#) and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

## Table of Contents

<a href="#">1.</a>	Introduction . . . . .	<a href="#">3</a>
<a href="#">2.</a>	Operational issues with default LFA tie breakers . . . . .	<a href="#">3</a>
<a href="#">2.1.</a>	Case 1: Edge router protecting core failures . . . . .	<a href="#">3</a>
2.2.	Case 2: Edge router chosen to protect core failures while core LFA exists . . . . .	<a href="#">5</a>
<a href="#">2.3.</a>	Case 3: suboptimal core alternate choice . . . . .	<a href="#">5</a>
<a href="#">2.4.</a>	Case 4: ISIS overload bit on LFA computing node . . . . .	<a href="#">6</a>
<a href="#">3.</a>	Need for coverage monitoring . . . . .	<a href="#">7</a>
<a href="#">4.</a>	Need for LFA activation granularity . . . . .	<a href="#">8</a>
<a href="#">5.</a>	Configuration requirements . . . . .	<a href="#">8</a>
<a href="#">5.1.</a>	LFA enabling/disabling scope . . . . .	<a href="#">8</a>
<a href="#">5.2.</a>	Policy based LFA selection . . . . .	<a href="#">9</a>
<a href="#">5.2.1.</a>	Connected vs remote alternates . . . . .	<a href="#">9</a>
<a href="#">5.2.2.</a>	Mandatory criteria . . . . .	<a href="#">10</a>
<a href="#">5.2.3.</a>	Enhanced criteria . . . . .	<a href="#">10</a>
<a href="#">5.2.4.</a>	Retrieving alternate path attributes . . . . .	<a href="#">11</a>
<a href="#">5.2.5.</a>	ECMP LFAs . . . . .	<a href="#">12</a>
<a href="#">5.2.6.</a>	SRLG . . . . .	<a href="#">13</a>
<a href="#">5.2.7.</a>	Link coloring . . . . .	<a href="#">14</a>
<a href="#">5.2.8.</a>	Bandwidth . . . . .	<a href="#">15</a>
<a href="#">5.2.9.</a>	Alternate preference . . . . .	<a href="#">16</a>
<a href="#">6.</a>	Operational aspects . . . . .	<a href="#">17</a>
<a href="#">6.1.</a>	ISIS overload bit on LFA computing node . . . . .	<a href="#">17</a>
<a href="#">6.2.</a>	Manual triggering of FRR . . . . .	<a href="#">18</a>

<a href="#">6.3.</a>	Required local information . . . . .	<a href="#">19</a>
<a href="#">6.4.</a>	Coverage monitoring . . . . .	<a href="#">19</a>
<a href="#">6.5.</a>	LFA and network planning . . . . .	<a href="#">20</a>
<a href="#">7.</a>	Security Considerations . . . . .	<a href="#">20</a>
<a href="#">8.</a>	Contributors . . . . .	<a href="#">20</a>

<a href="#">9.</a>	Acknowledgements . . . . .	<a href="#">21</a>
<a href="#">10.</a>	IANA Considerations . . . . .	<a href="#">21</a>
<a href="#">11.</a>	References . . . . .	<a href="#">21</a>
<a href="#">11.1.</a>	Normative References . . . . .	<a href="#">21</a>
<a href="#">11.2.</a>	Informative References . . . . .	<a href="#">21</a>
	Authors' Addresses . . . . .	<a href="#">22</a>

## [1.](#) Introduction

Following the first deployments of Loop Free Alternates (LFA), this document provides feedback to the community about the management of LFA.

[Section 2](#) provides real uses cases illustrating some limitations and suboptimal behavior.

[Section 4](#) proposes requirements for activation granularity and policy based selection of the alternate.

[Section 5](#) express requirements for the operational management of LFA.

## [2.](#) Operational issues with default LFA tie breakers

[RFC5286] introduces the notion of tie breakers when selecting the LFA among multiple candidate alternate next-hops. When multiple LFA exist, [RFC 5286](#) has favored the selection of the LFA providing the best coverage of the failure cases. While this is indeed a goal, this is one among multiple and in some deployment this lead to the selection of a suboptimal LFA. The following sections details real use cases of such limitations.

Note that the use case of per-prefix LFA is assumed throughout this analysis.

### [2.1.](#) Case 1: Edge router protecting core failures

Internet-Draft

LFA manageability

January 2015

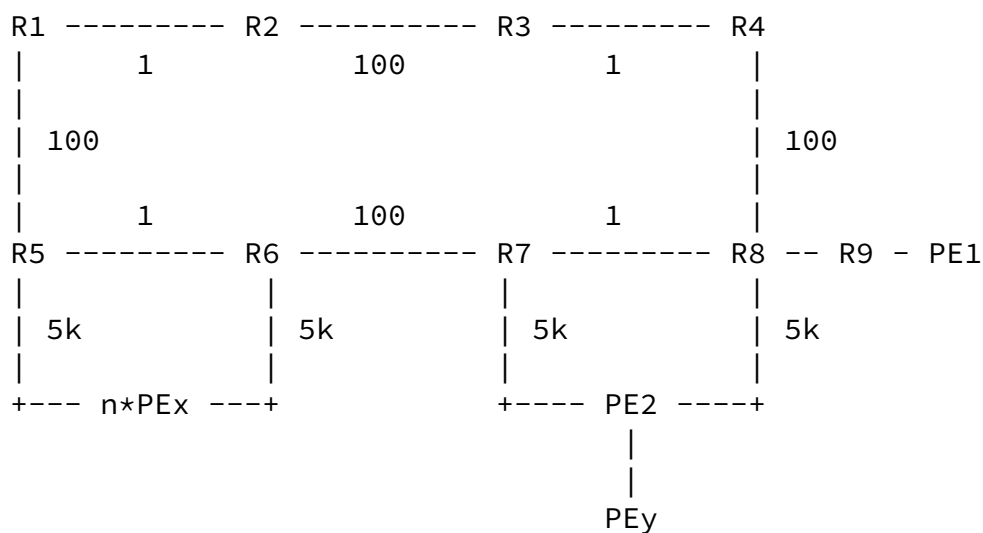


Figure 1

Rx routers are core routers using  $n \times 10G$  links. PEs are connected using links with lower bandwidth. PEx are a set of PEs connected to R5 and R6.

In figure 1, let us consider the traffic flowing from PE1 to PEx. The nominal path is R9-R8-R7-R6-PEx. Let us consider the failure of link R7-R8. For R8, R4 is not an LFA and the only available LFA is PE2.

When the core link R8-R7 fails, R8 switches all traffic destined to all the PEx towards the edge node PE2. Hence an edge node and edge links are used to protect the failure of a core link. Typically,

edge links have less capacity than core links and congestion may occur on PE2 links. Note that although PE2 was not directly affected by the failure, its links become congested and its traffic will suffer from the congestion.

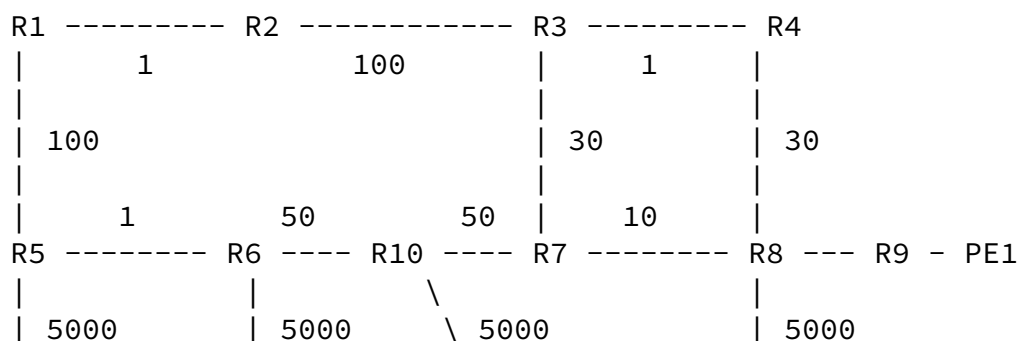
In summary, in case of failure, the impact on customer traffic is:

- o From PE2 point of view :
  - \* without LFA: no impact
  - \* with LFA: traffic is partially dropped (but possibly prioritized by a QoS mechanism). It must be highlighted that in such situation, traffic not affected by the failure may be affected by the congestion.
- o From R8 point of view:

- \* without LFA: traffic is totally dropped until convergence occurs.
- \* with LFA: traffic is partially dropped (but possibly prioritized by a QoS mechanism).

Besides the congestion aspects of using an Edge router as an alternate to protect a core failure, a service provider may consider this as a bad routing design and would like to prevent it.

## [2.2.](#) Case 2: Edge router choosen to protect core failures while core LFA exists



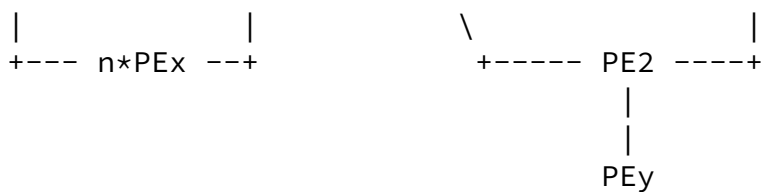


Figure 2

Rx routers are core routers meshed with  $n \times 10G$  links. PEs are meshed using links with lower bandwidth.

In the figure 2, let us consider the traffic coming from PE1 to PEx. Nominal path is R9-R8-R7-R10-R6-PEx. Let us consider the failure of the link R7-R8. For R8, R4 is a link-protecting LFA and PE2 is a node-protecting LFA. PE2 is chosen as best LFA due to its better protection type. Just like in case 1, this may lead to congestion on PE2 links upon LFA activation.

### [2.3.](#) Case 3: suboptimal core alternate choice

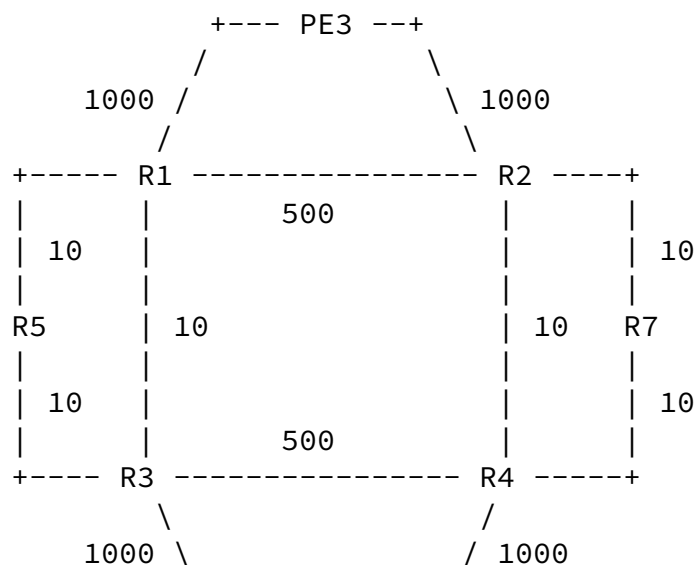




Figure 3

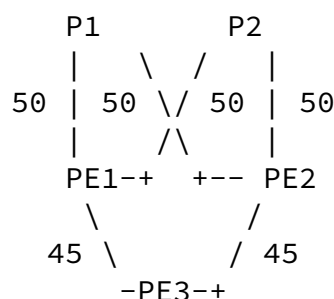
Rx routers are core routers. R1-R2 and R3-R4 links are 1G links. All others inter Rx links are 10G links.

In the figure above, let us consider the failure of link R1-R3. For destination PE3, R3 has two possible alternates:

- o R4, which is node-protecting
- o R5, which is link-protecting

R4 is chosen as best LFA due to its better protection type. However, it may not be desirable to use R4 for bandwidth capacity reason. A service provider may prefer to use high bandwidth links as preferred LFA. In this example, preferring shortest path over protection type may achieve the expected behavior, but in cases where metric are not reflecting bandwidth, it would not work and some other criteria would need to be involved when selecting the best LFA.

#### [2.4.](#) Case 4: ISIS overload bit on LFA computing node



(OL set)

Figure 4

In the figure above, PE3 has its overload bit set (permanently, for design reason) and wants to protect traffic using LFA for destination PE2.

On PE3, the loopfree condition is not satisfied :  $100 \nless 45 + 45$ . PE1 is thus not considered as an LFA. However thanks to the overload bit set on PE3, we know that PE1 is loopfree so PE1 is an LFA to reach PE2.

In case of overload condition set on a node, LFA behavior must be clarified.

### 3. Need for coverage monitoring

As per [[RFC6571](#)], LFA coverage highly depends on the used network topology. Even if remote LFA ([\[I-D.ietf-rtgwg-remote-lfa\]](#)) extends significantly the coverage of the basic LFA specification, there is still some cases where protection would not be available. As network topologies are constantly evolving (network extension, capacity addings, latency optimization ...), the protection coverage may change. Fast reroute functionality may be critical for some services supported by the network, a service provider must constantly know what protection coverage is currently available on the network. Moreover, predicting the protection coverage in case of network topology change is mandatory.

Today network simulation tool associated with whatif scenarios functionality are often used by service providers for the overall network design (capacity, path optimization ...). [Section 6.5](#), [Section 6.4](#) and [Section 6.3](#) of this document propose to add LFA informations into such tool and within routers, so a service provider may be able :

- o to evaluate protection coverage after a topology change.

- o to adjust the topology change to cover the primary need (e.g.



latency optimization or bandwidth increase) as well as LFA protection.

- o monitor constantly the LFA coverage in the live network and being alerted.

#### [4.](#) Need for LFA activation granularity

As all FRR mechanism, LFA installs backup paths in Forwarding Information Base (FIB). Depending of the hardware used by a service provider, FIB ressource may be critical. Activating LFA, by default, on all available components (IGP topologies, interface, address families ...) may lead to waste of FIB ressource as generally in a network only few destinations should be protected (e.g. loopback addresses supporting MPLS services) compared to the amount of destinations in RIB.

Moreover a service provider may implement multiple different FRR mechanism in its networks for different usages (MRT, TE FRR), computing LFAs for prefixes or interfaces that are already protected by another mechanism is useless.

[Section 5](#) of this document propose some implementation guidelines.

#### [5.](#) Configuration requirements

Controlling best alternate and LFA activation granularity is a requirement for Service Providers. This section defines configuration requirements for LFA.

##### [5.1.](#) LFA enabling/disabling scope

The granularity of LFA activation should be controlled (as alternate nexthop consume memory in forwarding plane).

An implementation of LFA SHOULD allow its activation with the following criteria:

- o Per address-family : ipv4 unicast, ipv6 unicast, LDP IPv4 unicast, LDP IPv6 unicast ...
- o Per routing context : VRF, virtual/logical router, global routing table, ...
- o Per interface
- o Per protocol instance, topology, area

- o Per prefixes: prefix protection SHOULD have a better priority compared to interface protection. This means that if a specific prefix must be protected due to a configuration request, LFA must be computed and installed for this prefix even if the primary outgoing interface is not configured for protection.

## [5.2.](#) Policy based LFA selection

When multiple alternates exist, LFA selection algorithm is based on tie breakers. Current tie breakers do not provide sufficient control on how the best alternate is chosen. This document proposes an enhanced tie breaker allowing service providers to manage all specific cases:

1. An implementation of LFA SHOULD support policy-based decision for determining the best LFA.
2. Policy based decision SHOULD be based on multiple criterions, with each criteria having a level of preference.
3. If the defined policy does not permit to determine a unique best LFA, an implementation SHOULD pick only one based on its own decision, as a default behavior. An implementation SHOULD also support election of multiple LFAs, for loadbalancing purposes.
4. Policy SHOULD be applicable to a protected interface or to a specific set of destinations. In case of application on the protected interface, all destinations primarily routed on this interface SHOULD use the interface policy.
5. It is an implementation choice to reevaluate policy dynamically or not (in case of policy change). If a dynamic approach is chosen, the implementation SHOULD recompute the best LFAs and reinstall them in FIB, without service disruption. If a non-dynamic approach is chosen, the policy would be taken into account upon the next IGP event. In this case, the implementation SHOULD support a command to manually force the recomputation/reinstallation of LFAs.

### [5.2.1.](#) Connected vs remote alternates

In addition to direct LFAs, tunnels (e.g. IP, LDP or RSVP-TE) to distant routers may be used to complement LFA coverage (tunnel tail used as virtual neighbor). When a router has multiple alternate candidates for a specific destination, it may have connected alternates and remote alternates reachable via a tunnel. Connected

alternates may not always provide an optimal routing path and it may be preferable to select a remote alternate over a connected

alternate. The usage of tunnels to extend LFA coverage is described in [[I-D.ietf-rtgwg-remote-lfa](#)].

In figure 1, there is no core alternate for R8 to reach PEs located behind R6, so R8 is using PE2 as alternate, which may generate congestion when FRR is activated. Instead, we could have a remote core alternate for R8 to protect PEs destinations. For example, a tunnel from R8 to R3 would ensure LFA protection without using an edge router to protect a core router.

When selecting the best alternate, the selection algorithm MUST consider all available alternates (connected or tunnel). Especially, computation of PQ set ([[I-D.ietf-rtgwg-remote-lfa](#)]) SHOULD be performed before best alternate selection.

#### [5.2.2.](#) Mandatory criteria

An implementation of LFA MUST support the following criteria:

- o Non candidate link: A link marked as "non candidate" will never be used as LFA.
- o A primary nexthop being protected by another primary nexthop of the same prefix (ECMP case).
- o Type of protection provided by the alternate: link protection, node protection. In case of node protection preference, an implementation SHOULD support fallback to link protection if node protection is not available.
- o Shortest path: lowest IGP metric used to reach the destination.
- o SRLG (as defined in [[RFC5286](#)] [Section 3](#), see also [Section 5.2.6](#) for more details).

#### [5.2.3.](#) Enhanced criteria

An implementation of LFA SHOULD support the following enhanced criteria:

- o Downstreamness of an alternate : preference of a downstream path over a non downstream path SHOULD be configurable.
- o Link coloring with : include, exclude and preference based system (see [Section 5.2.7](#)).
- o Link Bandwidth (see [Section 5.2.8](#)).

- o Alternate preference (see [Section 5.2.9](#)).

#### [5.2.4](#). Retrieving alternate path attributes

The policy to select the best alternate evaluate multiple criterions (e.g. metric, SRLG, link colors ...) which first need to be computed for each alternate.. In order to compare the different alternate path, a router must retrieve the attributes of each alternate path. The alternate path is composed of two distinct parts : PLR to alternate and alternate to destination.

##### [5.2.4.1](#). Connected alternate

For alternate path using a connected alternate :

- o attributes from PLR to alternate path are retrieved from the interface connected to the alternate.
- o attributes from alternate to destination path are retrieved from SPF rooted at the alternate. As the alternate is a connected alternate, the SPF has already been computed to find the alternate, so there is no need of additional computation.

##### [5.2.4.2](#). Remote alternate

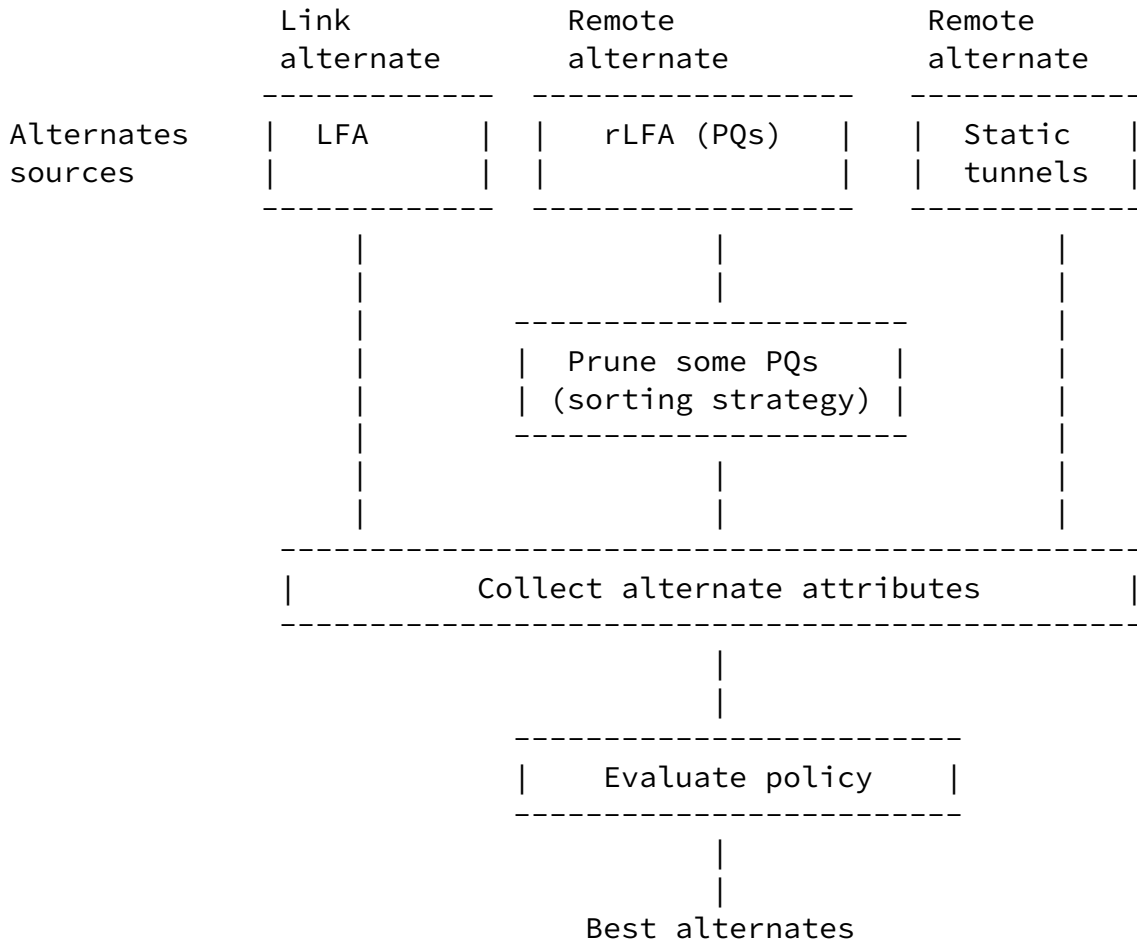
For alternate path using a remote alternate (tunnel) :

- o attributes from the PLR to alternate path are retrieved using the PLR's primary SPF if P space is used or using the neighbor's SPF if extended P space is used, combined with the attributes of the link(s) to reach that neighbor. In both cases, no additional SPF is required.

- o attributes from alternate to destination path are retrieved from SPF rooted at the remote alternate. An additional forward SPF is required for each remote alternate as indicated in [\[I-D.ietf-rtgwg-rlfa-node-protection\] section 3.2..](#)

The number of remote alternates may be very high, simulations shown that hundred's of PQs may exist for a single interface being protected. Running a forward SPF for every PQ-node in the network is not scalable.

To handle this situation, it is needed to limit the number of remote alternates to be evaluated to a finite number before collecting alternate path attributes and running the policy evaluation. [I-D.ietf-rtgwg-rlfa-node-protection] [Section 2.3.3](#) provides a way to reduce the number of PQ to be evaluated.



#### 5.2.5. ECMP LFAs

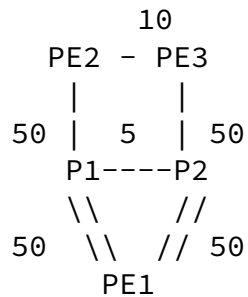


Figure 5

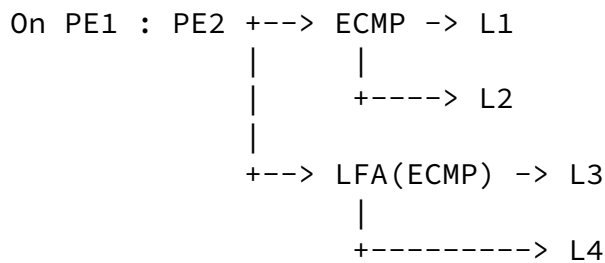
Links between P1 and PE1 are L1 and L2, links between P2 and PE1 are L3 and L4

In the figure above, primary path from PE1 to PE2 is through P1 using ECMP on two parallel links L1 and L2. In case of standard ECMP behavior, if L1 is failing, postconvergence nexthop would become L2 and there would be no longer ECMP. If LFA is activated, as stated in [\[RFC5286\] Section 3.4.](#), "alternate next-hops may themselves also be primary next-hops, but need not be" and "alternate next-hops should

maximize the coverage of the failure cases". In this scenario there is no alternate providing node protection, LFA will so prefer L2 as alternate to protect L1 which makes sense compared to postconvergence behavior.

Considering a different scenario using figure 5, where L1 and L2 are configured as a layer 3 bundle using a local feature, as well as L3/L4 being a second layer 3 bundle. Layer 3 bundles are configured as if a link in the bundle is failing, the traffic must be rerouted out of the bundle. Layer 3 bundles are generally introduced to increase bandwidth between nodes. In nominal situation, ECMP is still available from PE1 to PE2, but if L1 is failing, postconvergence nexthop would become ECMP on L3 and L4. In this case, LFA behavior SHOULD be adapted in order to reflect the bandwidth requirement.

We would expect the following FIB entry on PE1 :



If L1 or L2 is failing, traffic must be switched on the LFA ECMP bundle rather than using the other primary nexthop.

As mentioned in [\[RFC5286\] Section 3.4.](#), protecting a link within an ECMP by another primary nexthop is not a MUST. Moreover, we already presented in this document, that maximizing the coverage of the failure case may not be the right approach and policy based choice of alternate may be preferred.

An implementation SHOULD permit to prefer a primary nexthop by another primary nexthop with the possibility to deactivate this criteria. An implementation SHOULD permit to use an ECMP bundle as a LFA.

#### [5.2.6.](#) SRLG

[\[RFC5286\] Section 3.](#) proposes to reuse GMPLS IGP extensions to encode SRLGs ([\[RFC4205\]](#) and [\[RFC4203\]](#)). The section is also describing the algorithm to compute SRLG protection.

When SRLG protection is computed, and implementation SHOULD permit to :

- o Exclude alternates violating SRLG.
- o Maintain a preference system between alternates based on number of SRLG violations : more violations = less preference.

When applying SRLG criteria, the SRLG violation check SHOULD be performed on source to alternate as well as alternate to destination paths. In the case of remote LFA, PQ to destination path attributes

would be retrieved from SPT rooted at PQ.

#### 5.2.7. Link coloring

Link coloring is a powerful system to control the choice of alternates. Protecting interfaces are tagged with colors. Protected interfaces are configured to include some colors with a preference level, and exclude others.

Link color information SHOULD be signalled in the IGP. How signalling is done is out of scope of the document but it may be useful to reuse existing admin-groups from traffic-engineering extensions.

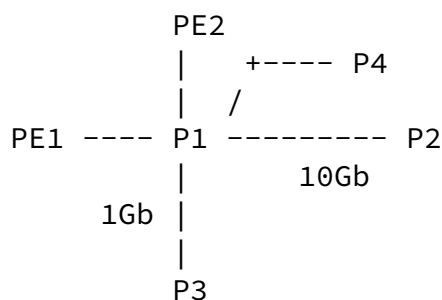


Figure 5

Example : P1 router is connected to three P routers and two PEs.

P1 is configured to protect the P1-P4 link. We assume that given the topology, all neighbors are candidate LFA. We would like to enforce a policy in the network where only a core router may protect against the failure of a core link, and where high capacity links are preferred.

In this example, we can use the proposed link coloring by:

- o Marking PEs links with color RED

- o Marking 10Gb CORE link with color BLUE
- o Marking 1Gb CORE link with color YELLOW



- o Configured the protected interface P1->P4 with :
  - \* Include BLUE, preference 200
  - \* Include YELLOW, preference 100
  - \* Exclude RED

Using this, PE links will never be used to protect against P1-P4 link failure and 10Gb link will be preferred.

The main advantage of this solution is that it can easily be duplicated on other interfaces and other nodes without change. A Service Provider has only to define the color system (associate color with a significance), as it is done already for TE affinities or BGP communities.

An implementation of link coloring:

- o SHOULD support multiple include and exclude colors on a single protected interface.
- o SHOULD provide a level of preference between included colors.
- o SHOULD support multiple colors configuration on a single protecting interface.

#### [5.2.8.](#) Bandwidth

As mentioned in previous sections, not taking into account bandwidth of an alternate could lead to congestion during FRR activation. We propose to base the bandwidth criteria on the link speed information for the following reason :

- o if a router S has a set of X destinations primarily forwarded to N, using per prefix LFA may lead to have a subset of X protected by a neighbor N1, another subset by N2, another subset by Nx ...
- o S is not aware about traffic flows to each destination and is not able to evaluate how much traffic will be sent to N1,N2, ... Nx in case of FRR activation.

Based on this, it is not useful to gather available bandwidth on alternate paths, as the router does not know how much bandwidth it

requires for protection. The proposed link speed approach provides a good approximation with a small cost as information is easily available.

The bandwidth criteria of the policy framework SHOULD work in two ways :

- o PRUNE : exclude a LFA if link speed to reach it is lower than the link speed of the primary nexthop interface.
- o PREFER : prefer a LFA based on his bandwidth to reach it compared to the link speed of the primary nexthop interface.

#### [5.2.9](#). Alternate preference

Rather than tagging interface on each node (using link color) to identify alternate node type (as example), it would be helpful if routers could be identified in the IGP. This would permit a grouped processing on multiple nodes. As an implementation need to exclude some specific alternates (see [Section 5.2.3](#)), an implementation :

- o SHOULD be able to give a preference to specific alternate.
- o SHOULD be able to give a preference to a group of alternate.
- o SHOULD be able to exclude a group of alternate.

A specific alternate may be identified by its interface, IP address or router ID and group of alternates may be identified by a marker (tag) (for example, in case of ISIS protocol [\[I-D.ietf-isis-node-admin-tag\]](#) can be used)

Consider the following network:

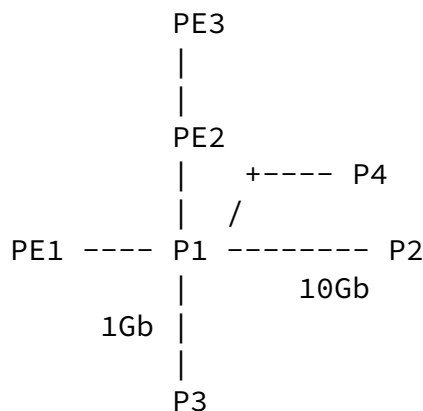


Figure 6

In the example above, each node is configured with a specific tag flooded through the IGP.

- o PE1,PE3: 200 (non candidate).
- o PE2: 100 (edge/core).
- o P1,P2,P3: 50 (core).

A simple policy could be configured on P1 to choose the best alternate for P1->P4 based on router function/role as follows :

- o criteria 1 -> alternate preference: exclude tag 100 and 200.
- o criteria 2 -> bandwidth.

## 6. Operational aspects

### 6.1. ISIS overload bit on LFA computing node

In [\[RFC5286\], Section 3.5](#), the setting of the overload bit condition in LFA computation is only taken into account for the case where a neighbor has the overload bit set.

In addition to [RFC 5286](#) inequality 1 Loop-Free Criterion ( $\text{Distance\_opt}(N, D) < \text{Distance\_opt}(N, S) + \text{Distance\_opt}(S, D)$ ), the IS-IS overload bit of the LFA calculating neighbor (S) SHOULD be taken into account. Indeed, if it has the overload bit set, no neighbor will loop back to traffic to itself.

## [6.2.](#) Manual triggering of FRR

Service providers often perform manual link shutdown (using router CLI) to perform some network changes/tests. A manual link shutdown may be done at multiple level : physical interface, logical interface, IGP interface, BFD session ... Especially testing or troubleshooting FRR requires to perform the manual shutdown on the remote end of the link as generally a local shutdown would not trigger FRR.

To enhance such situation, an implementation SHOULD support triggering/activating LFA Fast Reroute for a given link when a manual shutdown is done on a component that currently supports FRR activation.

An implementation MAY also support FRR activation for a specific interface or a specific prefix on a primary next-hop interface and revert without any action on any running component of the node (links or protocols). In this use case, the FRR activation time need to be controlled by a timer in case the operator forgot to revert traffic on primary path. When the timer expires, the traffic is automatically reverted to the primary path. This will make easier tests of fast-reroute path and then revert back to the primary path without causing a global network convergence.

For example :

- o if an implementation supports FRR activation upon BFD session down event, this implementation SHOULD support FRR activation when a manual shutdown is done on the BFD session. But if an implementation does not support FRR activation on BFD session down, there is no need for this implementation to support FRR activation on manual shutdown of BFD session.

- o if an implementation supports FRR activation on physical link down event (e.g. Rx laser Off detection, or error threshold raised ...), this implementation SHOULD support FRR activation when a manual shutdown at physical interface is done. But if an implementation does not support FRR activation on physical link down event, there is no need for this implementation to support FRR activation on manual physical link shutdown.
- o A CLI command may permit to switch from primary path to FRR path for testing FRR path for a specific. There is no impact on controlplane, only dataplane of the local node could be changed. A similar command may permit to switch back traffic from FRR path to primary path.

### [6.3.](#) Required local information

LFA introduction requires some enhancement in standard routing information provided by implementations. Moreover, due to the non 100% coverage, coverage informations is also required.

Hence an implementation :

- o MUST be able to display, for every prefixes, the primary nexthop as well as the alternate nexthop information.
- o MUST provide coverage information per activation domain of LFA (area, level, topology, instance, virtual router, address family ...).
- o MUST provide number of protected prefixes as well as non protected prefixes globally.
- o SHOULD provide number of protected prefixes as well as non protected prefixes per link.
- o MAY provide number of protected prefixes as well as non protected prefixes per priority if implementation supports prefix-priority insertion in RIB/FIB.
- o SHOULD provide a reason for chosing an alternate (policy and

criteria) and for excluding an alternate.

- o SHOULD provide the list of non protected prefixes and the reason why they are not protected (no protection required or no alternate available).

#### [6.4.](#) Coverage monitoring

It is pretty easy to evaluate the coverage of a network in a nominal situation, but topology changes may change the coverage. In some situations, the network may no longer be able to provide the required level of protection. Hence, it becomes very important for service providers to get alerted about changes of coverage.

An implementation SHOULD :

- o provide an alert system if total coverage (for a node) is below a defined threshold or comes back to a normal situation.
- o provide an alert system if coverage of a specific link is below a defined threshold or comes back to a normal situation.

An implementation MAY :

- o provide an alert system if a specific destination is not protected anymore or when protection comes back up for this destination

Although the procedures for providing alerts are beyond the scope of this document, we recommend that implementations consider standard and well used mechanisms like syslog or SNMP traps.

#### [6.5.](#) LFA and network planning

The operator may choose to run simulations in order to ensure full coverage of a certain type for the whole network or a given subset of the network. This is particularly likely if he operates the network in the sense of the third backbone profiles described in [[RFC6571](#)], that is, he seeks to design and engineer the network topology in a way that a certain coverage is always achieved. Obviously a complete and exact simulation of the IP FRR coverage can only be achieved, if the behavior is deterministic and if the algorithm used is available

to the simulation tool. Thus, an implementation SHOULD:

- o Behave deterministic in its selection LFA process. I.e. in the same topology and with the same policy configuration, the implementation MUST always choose the same alternate for a given prefix.
- o Document its behavior. The implementation SHOULD provide enough documentation of its behavior that allows an implementer of a simulation tool, to foresee the exact choice of the LFA implementation for every prefix in a given topology. This SHOULD take into account all possible policy configuration options. One possible way to document this behavior is to disclose the algorithm used to choose alternates.

## 7. Security Considerations

This document does not introduce any change in security consideration compared to [[RFC5286](#)].

## 8. Contributors

Significant contributions were made by Pierre Francois, Hannes Gredler, Chris Bowers, Jeff Tantsura, Uma Chunduri and Mustapha Aissaoui which the authors would like to acknowledge.

## 9. Acknowledgements

## 10. IANA Considerations

This document has no action for IANA.

## 11. References

### 11.1. Normative References

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

- [RFC4203] Kompella, K. and Y. Rekhter, "OSPF Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", [RFC 4203](#), October 2005.
- [RFC4205] Kompella, K. and Y. Rekhter, "Intermediate System to Intermediate System (IS-IS) Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", [RFC 4205](#), October 2005.
- [RFC5286] Atlas, A. and A. Zinin, "Basic Specification for IP Fast Reroute: Loop-Free Alternates", [RFC 5286](#), September 2008.
- [RFC5307] Kompella, K. and Y. Rekhter, "IS-IS Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", [RFC 5307](#), October 2008.
- [RFC6571] Filsfils, C., Francois, P., Shand, M., Decraene, B., Uttaro, J., Leymann, N., and M. Horneffer, "Loop-Free Alternate (LFA) Applicability in Service Provider (SP) Networks", [RFC 6571](#), June 2012.

## [11.2.](#) Informative References

- [I-D.ietf-isis-node-admin-tag]  
Sarkar, P., Gredler, H., Hegde, S., Litkowski, S., Decraene, B., Li, Z., Aries, E., Rodriguez, R., and H. Raghuvver, "Advertising Per-node Admin Tags in IS-IS", [draft-ietf-isis-node-admin-tag-00](#) (work in progress), December 2014.
- [I-D.ietf-rtgwg-remote-lfa]  
Bryant, S., Filsfils, C., Previdi, S., Shand, M., and N. So, "Remote Loop-Free Alternate Fast Re-Route", [draft-ietf-rtgwg-remote-lfa-10](#) (work in progress), January 2015.

- [I-D.ietf-rtgwg-rlfa-node-protection]  
Sarkar, P., Gredler, H., Hegde, S., Bowers, C., Litkowski, S., and H. Raghuvver, "Remote-LFA Node Protection and Manageability", [draft-ietf-rtgwg-rlfa-node-protection-01](#) (work in progress), December 2014.



## Authors' Addresses

Stephane Litkowski  
Orange

Email: [stephane.litkowski@orange.com](mailto:stephane.litkowski@orange.com)

Bruno Decraene  
Orange

Email: [bruno.decraene@orange.com](mailto:bruno.decraene@orange.com)

Clarence Filsfils  
Cisco Systems

Email: [cfilsfil@cisco.com](mailto:cfilsfil@cisco.com)

Kamran Raza  
Cisco Systems

Email: [skraza@cisco.com](mailto:skraza@cisco.com)

Martin Horneffer  
Deutsche Telekom

Email: [Martin.Horneffer@telekom.de](mailto:Martin.Horneffer@telekom.de)

Pushpasis Sarkar  
Juniper Networks

Email: [psarkar@juniper.net](mailto:psarkar@juniper.net)