Internet Engineering Task Force (IETF)
Request for Comments: 8333
Category: Standards Track
ISSN: 2070-1721

S. Litkowski
B. Decraene
Orange
C. Filsfils
Cisco Systems
P. Francois
Individual Contributor

March 2018

Micro-loop Prevention by Introducing a Local Convergence Delay

Abstract

This document describes a mechanism for link-state routing protocols that prevents local transient forwarding loops in case of link failure. This mechanism proposes a two-step convergence by introducing a delay between the convergence of the node adjacent to the topology change and the network-wide convergence.

Because this mechanism delays the IGP convergence, it may only be used for planned maintenance or when Fast Reroute (FRR) protects the traffic during the time between the link failure and the IGP convergence.

The mechanism is limited to the link-down event in order to keep the mechanism simple.

Simulations using real network topologies have been performed and show that local loops are a significant portion (>50%) of the total forwarding loops.

Status of This Memo

Table of Contents

1. Introduction

Micro-loops and some potential solutions are described in [RFC5715]. This document describes a simple targeted mechanism that prevents micro-loops that are local to the failure. Based on network analysis, local micro-loops make up a significant portion of the micro-loops. A simple and easily deployable solution for these local micro-loops is critical because these local loops cause some traffic loss after an FRR alternate has been used (see Section 3.1).

Consider the case in Figure 1 where S does not have an LFA (Loop-Free Alternate) to protect its traffic to D when the S-D link fails. That means that all non-D neighbors of S on the topology will send to S any traffic destined to D; if a neighbor did not, then that neighbor would be loop-free. Regardless of the advanced FRR technique used, when S converges to the new topology, it will send its traffic to a neighbor that is not loop-free and will thus cause a local micro-loop. The deployment of advanced FRR techniques motivates this simple router-local mechanism to solve this targeted problem. This solution can work with the various techniques described in [RFC5715].

```
D ------ C
|        |
|        | 5
|        |
S ------ B
```

Figure 1

In Figure 1, all links have a metric of 1 except the B-C link, which has a metric of 5. When the S-D link fails, a transient forwarding loop may appear between S and B if S updates its forwarding entry to D before B does.

2. Terminology

2.1. Acronyms

FIB: Forwarding Information Base

FRR: Fast Reroute

IGP: Interior Gateway Protocol

LFA: Loop-Free Alternate

LSA: Link State Advertisement

LSP: Link State Packet

MRT: Maximally Redundant Tree

oFIB: Ordered FIB

PLR: Point of Local Repair

PLSN: Path Locking via Safe Neighbors

RIB: Routing Information Base

RLFA: Remote Loop-Free Alternate

SPF: Shortest Path First

TTL: Time to Live

2.2. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

3. Side Effects of Transient Forwarding Loops

Even if they are very limited in duration, transient forwarding loops may cause significant network damage.

3.1. FRR Inefficiency

In Figure 2, we consider an IP/LDP routed network.

```
                                 D
                               1 |
                                 |    1
                                 A ------ B
                                 |        |    ^
                              10 |        | 5  | T
                                 |        |    |
                                 E--------C
                                 |    1
                               1 |
                                 S

                                 Figure 2
```

An RSVP-TE tunnel T, provisioned on C and terminating on B, is used to protect the traffic against C-B link failure (the IGP shortcut feature, defined in [RFC3906], is activated on C). The primary path of T is C->B and FRR is activated on T, providing an FRR bypass or detour using path C->E->A->B. On router C, the next hop to D is the tunnel T, thanks to the IGP shortcut. When the C-B link fails:

1. C detects the failure and updates the tunnel path using a preprogrammed FRR path. The traffic path from S to D becomes S->E->C->E->A->B->A->D.

2. In parallel, on router C, both the IGP convergence and the TE tunnel convergence (tunnel path recomputation) are occurring:

   * The tunnel T path is recomputed and now uses C->E->A->B.

   * The IGP path to D is recomputed and now uses C->E->A->D.

3. On C, the tail-end of the TE tunnel (router B) is no longer on the shortest-path tree (SPT) to D, so C does not continue to encapsulate the traffic to D using the tunnel T and updates its forwarding entry to D using the next-hop E.

If C updates its forwarding entry to D before router E, there would be a transient forwarding loop between C and E until E has converged.

Table 1 describes a theoretical sequence of events happening when the B-C link fails. This theoretical sequence of events should only be read as an example.

Network Condition	Time	Router C Events	Router E Events
S->D Traffic OK			
S->D Traffic lost	t0	Link B-C fails	Link B-C fails
	t0+20 ms	C detects the failure	

S->D Traffic OK	t0+40 ms	C activates FRR	
	t0+50 ms	C updates its local LSP/LSA	
	t0+60 ms	C floods its local updated LSP/LSA	
	t0+62 ms	C schedules SPF (100 ms)	
	t0+87 ms		E receives LSP/LSA from C and floods it
	t0+92 ms		E schedules SPF (100 ms)
	t0+163 ms	C computes SPF	
	t0+165 ms	C starts updating its RIB/FIB	
	t0+193 ms		E computes SPF
	t0+199 ms		E starts updating its RIB/FIB
S->D Traffic lost	t0+255 ms	C updates its RIB/FIB for D	
	t0+340 ms	C convergence ends	
S->D Traffic OK	t0+443 ms		E updates its RIB/FIB for D
	t0+470 ms		E convergence ends

Table 1

The issue described here is completely independent of the FRR mechanism involved (e.g., TE FRR, LFA/RLFA, MRT, etc.) when the primary path uses hop-by-hop routing. The protection enabled by FRR works perfectly but only ensures protection until the PLR has converged (as soon as the PLR has converged, it replaces its FRR path with a new primary path). When implementing FRR, a service provider wants to guarantee a very limited loss of connectivity time. The example described in this section shows that the benefit of FRR may be completely lost due to a transient forwarding loop appearing when PLR has converged. Delaying FIB updates after the IGP convergence (1) may allow the FRR path to be kept until the neighbors have converged and (2) preserves the customer traffic.

3.2. Network Congestion

In Figure 3, when the S-D link fails, a transient forwarding loop may appear between S and B for destination D. The traffic on the S-B link will constantly increase due to the looping traffic to D. Depending on the TTL of the packets, the traffic rate destined to D, and the bandwidth of the link, the S-B link may become congested in a few hundreds of milliseconds and will stay congested until the loop is eliminated.

```
                      1
                  D ------ C
                  |        |
                1 |        | 5
                  |        |
             A -- S ------ B
                / |    1
               F  E
```

Figure 3

The congestion introduced by transient forwarding loops is problematic as it can affect traffic that is not directly affected by the failing network component. In Figure 3, the congestion of the S-B link will impact some customer traffic that is not directly affected by the failure, e.g., traffic from A to B, F to B, and E to B. Class of service may mitigate the congestion for some traffic. However, some traffic not directly affected by the failure will still be dropped as a router is not able to distinguish the looping traffic from the normally forwarded traffic.

4. Overview of the Solution

This document defines a two-step convergence initiated by the router detecting a failure and advertising the topological change in the IGP. This introduces a delay between network-wide convergence and the convergence of the local router.

The solution described in this document is limited to local link-down events in order to keep the solution simple.

This ordered convergence is similar to the ordered FIB (oFIB) approach defined in [RFC6976], but it is limited to only a "one-hop" distance. As a consequence, it is more simple and becomes a local-only feature that does not require interoperability. This benefit comes with the limitation of eliminating transient forwarding loops involving the local router only. The mechanism also reuses some concepts described in [PLSN].

5. Specification

5.1. Definitions

This document refers to the following existing IGP timers. These timers may be standardized or implemented as a vendor-specific local feature.

- o LSP_GEN_TIMER: The delay between the consecutive generation of two local LSPs/LSAs. From an operational point of view, this delay is usually tuned to batch multiple local events in a single local LSP/LSA update. In IS-IS, this timer is defined as minimumLSPGenerationInterval [ISO10589]. In OSPF version 2, this timer is defined as MinLSInterval [RFC2328]. It is often associated with a vendor-specific damping mechanism to slow down reactions by incrementing the timer when multiple consecutive events are detected.

- o SPF_DELAY: The delay between the first IGP event triggering a new routing table computation and the start of that routing table computation. It is often associated with a damping mechanism to slow down reactions by incrementing the timer when the IGP becomes unstable. As an example, [BACKOFF] defines a standard SPF delay algorithm.

This document introduces the following new timer:

- o ULOOP_DELAY_DOWN_TIMER: Used to slow down the local node convergence in case of link-down events.

5.2. Regular IGP Reaction

When the status of an adjacency or link changes, the regular IGP convergence behavior of the router advertising the event involves the following main steps:

1. IGP is notified of the up/down event.

2. The IGP processes the notification and postpones the reaction for LSP_GEN_TIMER ms.

3. Upon LSP_GEN_TIMER expiration, the IGP updates its LSP/LSA and floods it.

4. The SPF computation is scheduled in SPF_DELAY ms.

5. Upon SPF_DELAY timer expiration, the SPF is computed, and then the RIB and FIB are updated.

5.3. Local Events

The mechanism described in this document assumes that there has been a single link failure as seen by the IGP area/level. If this assumption is violated (e.g., multiple links or nodes failed), then regular IP convergence must be applied (as described in Section 5.2).

To determine if the mechanism is applicable or not, an implementation SHOULD implement logic to correlate the protocol messages (LSP/LSA) received during the SPF scheduling period in order to determine the topology changes that occurred. This is necessary as multiple protocol messages may describe the same topology change, and a single protocol message may describe multiple topology changes. As a consequence, determining a particular topology change MUST be independent of the order of reception of those protocol messages. How the logic works is left to the implementation.

Using this logic, if an implementation determines that the associated topology change is a single local link failure, then the router MAY use the mechanism described in this document; otherwise, the regular IP convergence MUST be used.

In Figure 4, let router B be the computing router when the link B-C fails. B updates its local LSP/LSA describing the link B-C as down, C does the same, and both start flooding their updated LSPs/LSAs. During the SPF_DELAY period, B and C learn all the LSPs/LSAs to consider. B sees that C is flooding an advertisement that indicates that a link is down, and B is the other end of that link. B determines that B and C are describing the same single event. Since B receives no other changes, B can determine that this is a local link failure and may decide to activate the mechanism described in this document.

```
              +--- E ----+--------+
              |          |        |
       A ---- B -------- C ------ D
```

Figure 4

5.4. Local Delay for Link-Down Events

This document introduces a change in step 5 (see list in Section 5.2) so that, upon an adjacency or link-down event, the local convergence is delayed compared to the network-wide convergence. The new step 5 is described below:

5. Upon SPF_DELAY timer expiration, the SPF is computed. If the condition of a single local link-down event has been met, then an update of the RIB and the FIB MUST be delayed for ULOOP_DELAY_DOWN_TIMER ms. Otherwise, the RIB and FIB SHOULD be updated immediately.

If a new convergence occurs while ULOOP_DELAY_DOWN_TIMER is running, ULOOP_DELAY_DOWN_TIMER is stopped, and the RIB/FIB SHOULD be updated as part of the new convergence event.

As a result of this addition, routers local to the failure will converge slower than remote routers. Hence, it SHOULD only be done for a non-urgent convergence, such as administrative deactivation (maintenance) or when the traffic is protected by FRR.

6. Applicability

As previously stated, this mechanism only avoids the forwarding loops on the links between the node local to the failure and its neighbors. Forwarding loops may still occur on other links.

6.1. Applicable Case: Local Loops

In Figure 5, let us consider the traffic from G to F. The primary path is G->D->C->E->F. When the link C-E fails, if C updates its forwarding entry for F before D, a transient loop occurs. This is sub-optimal as it breaks C's FRR forwarding even though upstream routers are still forwarding the traffic to C.

```
                          A ------ B ----- E
                          |              / |
                          |             /  |
                      G---D------------C   F

                      All the links have a metric of 1
```

Figure 5

By implementing the mechanism defined in this document on C, when the C-E link fails, C delays the update of its forwarding entry to F, in order to allow some time for D to converge. FRR on C keeps protecting the traffic during this period. When ULOOP_DELAY_DOWN_TIMER expires on C, its forwarding entry to F is updated. There is no transient forwarding loop on the link C-D.

6.2. Non-applicable Case: Remote Loops

In Figure 6, let us consider the traffic from G to K. The primary path is G->D->C->F->J->K. When the C-F link fails, if C updates its forwarding entry to K before D, a transient loop occurs between C and D.

```
                   A ------ B ----- E --- H
                   |                      |
                   |                      |
               G---D--------C ------F --- J ---- K

               All the links have a metric of 1 except B-E=15
```

Figure 6

By implementing the mechanism defined in this document on C, when the link C-F fails, C delays the update of its forwarding entry to K, allowing time for D to converge. When ULOOP_DELAY_DOWN_TIMER expires on C, its forwarding entry to F is updated. There is no transient forwarding loop between C and D. However, a transient forwarding loop may still occur between D and A. In this scenario, this mechanism is not enough to address all the possible forwarding loops. However, it does not create additional traffic loss. Besides, in some cases -- such as when the nodes update their FIB in the order C, A, D because the router A is quicker than D to converge -- the mechanism may still avoid the forwarding loop that would have otherwise occurred.

7. Simulations

Simulations have been run on multiple service-provider topologies. We evaluated the efficiency of the mechanism on eight different service-provider topologies (different network size and design). Table 2 displays the gain for each topology.

Topology	Gain
T1	71%
T2	81%
T3	62%
T4	50%
T5	70%
T6	70%
T7	59%
T8	77%

Table 2

We evaluated the gain as follows:

- o We considered a tuple (link A-B, destination D, PLR S, backup next-hop N) as a loop if, upon link A-B failure, the flow from a router S upstream from A (A could be considered as PLR also) to D may loop due to convergence time difference between S and one of its neighbors N.

- o We evaluated the number of potential loop tuples in normal conditions.

- o We evaluated the number of potential loop tuples using the same topological input but taking into account that S converges after N.

- o The gain is the relative number of loops (both remote and local) we succeed in suppressing.

For topology 1, implementing the local delay prevented 71% of the transient forwarding loops created by the failure of any link. The analysis shows that all local loops are prevented and only remote loops remain.

8. Deployment Considerations

Transient forwarding loops have the following drawbacks:

- They limit FRR efficiency. Even if FRR is activated within 50 ms, as soon as the PLR has converged, the traffic may be affected by a transient loop.

- They may impact traffic not directly affected by the failure (due to link congestion).

The local delay mechanism is a transient forwarding loop avoidance mechanism (like oFIB). Even if it only addresses local transient loops, the efficiency versus complexity comparison of the mechanism makes it a good solution. It is also incrementally deployable with incremental benefits, which makes it an attractive option for both vendors to implement and service providers to deploy. Delaying the convergence time is not an issue if we consider that the traffic is protected during the convergence.

The ULOOP_DELAY_DOWN_TIMER value should be set according to the maximum IGP convergence time observed in the network (usually observed in the slowest node).

This mechanism is limited to link-down events. When a link goes down, it eventually goes back up. As a consequence, with this mechanism deployed, only the link-down event will be protected against transient forwarding loops while the link-up event will not. If the operator wants to limit the impact of transient forwarding loops during the link-up event, it should make sure to use specific procedures to bring the link back online. As examples, the operator can decide to put the link back online outside of business hours, or it can use some incremental metric changes to prevent loops (as proposed in [RFC5715]).

9. Examples

We consider the following figure for the examples in this section:

```
                              D
                            1 |        F----X
                              |    1   |
                              A ------ B
                              |        |
                           10 |        | 5
                              |        |
                              E--------C
                              |    1
                            1 |
                              S
```

Figure 7

The network above is considered to have a convergence time of about 1 second, so ULOOP_DELAY_DOWN_TIMER will be adjusted to this value. We also consider that FRR is running on each node.

9.1. Local Link-Down Event

Table 3 describes the events and their timing on routers C and E when the link B-C goes down. It is based on a theoretical sequence of events that should only been read as an example. As C detects a single local event corresponding to a link-down event (its LSP + LSP from B received), it applies the local delay down behavior, and no micro-loop is formed.

Network Condition	Time	Router C Events	Router E Events
S->D Traffic OK			
S->D Traffic lost	t0	Link B-C fails	Link B-C fails
	t0+20 ms	C detects the failure	
S->D Traffic OK	t0+40 ms	C activates FRR	
	t0+50 ms	C updates its local LSP/LSA	
	t0+53 ms	C floods its local updated LSP/LSA	
	t0+60 ms	C schedules SPF (100 ms)	
	t0+67 ms	C receives LSP/LSA from B and floods it	
	t0+87 ms		E receives LSP/LSA from C and floods it
	t0+90 ms		E schedules SPF (100 ms)
	t0+161 ms	C computes SPF	
	t0+165 ms	C delays its RIB/FIB update (1 sec)	
	t0+193 ms		E computes SPF
	t0+199 ms		E starts updating its RIB/FIB

	t0+443 ms		E updates its RIB/FIB for D
	t0+470 ms		E convergence ends
	t0+1165 ms	C starts updating its RIB/FIB	
	t0+1255 ms	C updates its RIB/FIB for D	
	t0+1340 ms	C convergence ends	

Table 3

Similarly, upon B-C link-down event, if LSP/LSA from B is received before C detects the link failure, C will apply the route update delay if the local detection is part of the same SPF run. Table 4 describes the associated theoretical sequence of events. It should only been read as an example.

Network Condition	Time	Router C Events	Router E Events
S->D Traffic OK			
S->D Traffic lost	t0	Link B-C fails	Link B-C fails
	t0+32 ms	C receives LSP/LSA from B and floods it	
	t0+33 ms	C schedules SPF (100 ms)	
	t0+50 ms	C detects the failure	

S->D Traffic OK	t0+55 ms	C activates FRR	
	t0+55 ms	C updates its local LSP/LSA	
	t0+70 ms	C floods its local updated LSP/LSA	
	t0+87 ms		E receives LSP/LSA from C and floods it
	t0+90 ms		E schedules SPF (100 ms)
	t0+135 ms	C computes SPF	
	t0+140 ms	C delays its RIB/FIB update (1 sec)	
	t0+193 ms		E computes SPF
	t0+199 ms		E starts updating its RIB/FIB
	t0+443 ms		E updates its RIB/FIB for D
	t0+470 ms		E convergence ends
	t0+1145 ms	C starts updating its RIB/FIB	
	t0+1255 ms	C updates its RIB/FIB for D	
	t0+1340 ms	C convergence ends	

Table 4

9.2. Local and Remote Event

Table 5 describes the events and their timing on router C and E when the link B-C goes down and when the link F-X fails in the same time window. C will not apply the local delay because a non-local topology change is also received. Table 5 is based on a theoretical sequence of events that should only been read as an example.

Network Condition	Time	Router C Events	Router E Events
S->D Traffic OK			
S->D Traffic lost	t0	Link B-C fails	Link B-C fails
	t0+20 ms	C detects the failure	
	t0+36 ms	Link F-X fails	Link F-X fails
S->D Traffic OK	t0+40 ms	C activates FRR	
	t0+50 ms	C updates its local LSP/LSA	
	t0+54 ms	C receives LSP/LSA from F and floods it	
	t0+60 ms	C schedules SPF (100 ms)	
	t0+67 ms	C receives LSP/LSA from B and floods it	
	t0+69 ms		E receives LSP/LSA from F, floods it and schedules SPF (100 ms)

	t0+70 ms	C floods its local updated LSP/LSA	
	t0+87 ms		E receives LSP/LSA from C
	t0+117 ms		E floods LSP/LSA from C
	t0+160 ms	C computes SPF	
	t0+165 ms	C starts updating its RIB/FIB (NO DELAY)	
	t0+170 ms		E computes SPF
	t0+173 ms		E starts updating its RIB/FIB
S->D Traffic lost	t0+365 ms	C updates its RIB/FIB for D	
S->D Traffic OK	t0+443 ms		E updates its RIB/FIB for D
	t0+450 ms	C convergence ends	
	t0+470 ms		E convergence ends

Table 5

9.3. Aborting Local Delay

Table 6 describes the events and their timing on routers C and E when the link B-C goes down. In addition, we consider what happens when the F-X link fails during local delay of the FIB update. C will first apply the local delay, but when the new event happens, it will fall back to the standard convergence mechanism without further delaying route insertion. In this example, we consider a ULOOP_DELAY_DOWN_TIMER configured to 2 seconds. Table 6 is based on a theoretical sequence of events that should only been read as an example.

Network Condition	Time	Router C Events	Router E Events
S->D Traffic OK			
S->D Traffic lost	t0	Link B-C fails	Link B-C fails
	t0+20 ms	C detects the failure	
S->D Traffic OK	t0+40 ms	C activates FRR	
	t0+50 ms	C updates its local LSP/LSA	
	t0+55 ms	C floods its local updated LSP/LSA	
	t0+57 ms	C schedules SPF (100 ms)	
	t0+67 ms	C receives LSP/LSA from B and floods it	
	t0+87 ms		E receives LSP/LSA from C and floods it
	t0+90 ms		E schedules SPF (100 ms)

```
   |            | t0+160 |    C computes SPF    |                      |
   |            |   ms   |                      |                      |
   |            |        |                      |                      |
   |            | t0+165 | C delays its RIB/FIB |                      |
   |            |   ms   |    update (2 sec)    |                      |
   |            |        |                      |                      |
   |            | t0+193 |                      |    E computes SPF    |
   |            |   ms   |                      |                      |
   |            |        |                      |                      |
   |            | t0+199 |                      |  E starts updating   |
   |            |   ms   |                      |     its RIB/FIB      |
   |            |        |                      |                      |
   |            | t0+254 |    Link F-X fails    |    Link F-X fails    |
   |            |   ms   |                      |                      |
   |            |        |                      |                      |
   |            | t0+300 |  C receives LSP/LSA  |                      |
   |            |   ms   | from F and floods it |                      |
   |            |        |                      |                      |
   |            | t0+303 | C schedules SPF (200 |                      |
   |            |   ms   |         ms)          |                      |
   |            |        |                      |                      |
   |            | t0+312 |  E receives LSP/LSA  |                      |
   |            |   ms   | from F and floods it |                      |
   |            |        |                      |                      |
   |            | t0+313 | E schedules SPF (200 |                      |
   |            |   ms   |         ms)          |                      |
   |            |        |                      |                      |
   |            | t0+502 |    C computes SPF    |                      |
   |            |   ms   |                      |                      |
   |            |        |                      |                      |
   |            | t0+505 |  C starts updating   |                      |
   |            |   ms   |   its RIB/FIB (NO    |                      |
   |            |        |        DELAY)        |                      |
   |            |        |                      |                      |
   |            | t0+514 |                      |    E computes SPF    |
   |            |   ms   |                      |                      |
   |            |        |                      |                      |
   |            | t0+519 |                      |  E starts updating   |
   |            |   ms   |                      |     its RIB/FIB      |
   |            |        |                      |                      |
   |    S->D    | t0+659 |    C updates its     |                      |
   |  Traffic   |   ms   |    RIB/FIB for D     |                      |
   |    lost    |        |                      |                      |
   |            |        |                      |                      |
```

S->D Traffic OK	t0+778 ms		E updates its RIB/FIB for D
	t0+781 ms	C convergence ends	
	t0+810 ms		E convergence ends

Table 6

10. Comparison with Other Solutions

As stated in Section 4, the local delay solution reuses some concepts already introduced by other IETF proposals but tries to find a trade-off between efficiency and simplicity. This section tries to compare behaviors of the solutions.

10.1. PLSN

PLSN [PLSN] describes a mechanism where each node in the network tries to avoid transient forwarding loops upon a topology change by always keeping traffic on a loop-free path for a defined duration (locked path to a safe neighbor). The locked path may be the new primary next hop, another neighbor, or the old primary next hop depending on how the safety condition is satisfied.

PLSN does not solve all transient forwarding loops (see Section 4 of [PLSN] for more details).

The solution defined in this document reuses some concepts of PLSN but in a more simple fashion:

o PLSN has three different behaviors: (1) keep using the old next hop, (2) use the new primary next hop if it is safe, or (3) use another safe next hop. The local delay solution, however, only has one: keep using the current next hop (i.e., the old primary next hop or an already-activated FRR path).

o PLSN may cause some damage while using a safe next hop that is not the new primary next hop if the new safe next hop does not provide enough bandwidth (see [RFC7916]). The solution defined in this document may not experience this issue as the service provider may have control on the FRR path being used, preventing network congestion.

- o PLSN applies to all nodes in a network (remote or local changes), while the mechanism defined in this document applies only to the nodes connected to the topology change.

10.2. oFIB

oFIB [RFC6976] describes a mechanism where the convergence of the network upon a topology change is ordered in order to prevent transient forwarding loops. Each router in the network deduces the failure type from the LSA/LSP received and computes/applies a specific FIB update timer based on the failure type and its rank in the network, considering the failure point as root.

The oFIB mechanism solves all the transient forwarding loops in a network at the price of introducing complexity in the convergence process that may require careful monitoring by the service provider.

The solution defined in this document reuses the oFIB concept but limits it to the first hop that experiences the topology change. As demonstrated, the mechanism defined in this document allows all the local transient forwarding loops to be solved; these represent a high percentage of all the loops. Moreover, limiting to one hop allows network-wide convergence behavior to be kept.

11. IANA Considerations

This document has no IANA actions.

12. Security Considerations

This document does not introduce any change in terms of IGP security. The operation is internal to the router. The local delay does not increase the number of attack vectors as an attacker could only trigger this mechanism if it already has the ability to disable or enable an IGP link. The local delay does not increase the negative consequences. If an attacker has the ability to disable or enable an IGP link, it can already harm the network by creating instability and harm the traffic by creating forwarding packet loss and forwarding loss for the traffic crossing that link.

13. References

13.1. Normative References

[ISO10589] International Organization for Standardization, "Information technology -- Telecommunications and information exchange between systems -- Intermediate System to Intermediate System intra-domain routeing information exchange protocol for use in conjunction with the protocol for providing the connectionless-mode network service (ISO 8473)", ISO/IEC 10589:2002, Second Edition, November 2002.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <https://www.rfc-editor.org/info/rfc2119>.

[RFC2328] Moy, J., "OSPF Version 2", STD 54, RFC 2328, DOI 10.17487/RFC2328, April 1998, <https://www.rfc-editor.org/info/rfc2328>.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <https://www.rfc-editor.org/info/rfc8174>.

13.2. Informative References

[BACKOFF] Decraene, B., Litkowski, S., Gredler, H., Lindem, A., Francois, P., and C. Bowers, "SPF Back-off Delay algorithm for link state IGPs", Work in Progress, draft-ietf-rtgwg-backoff-algo-10, March 2018.

[PLSN] Zinin, A., "Analysis and Minimization of Microloops in Link-state Routing Protocols", Work in Progress, draft-ietf-rtgwg-microloop-analysis-01, October 2005.

[RFC3906] Shen, N. and H. Smit, "Calculating Interior Gateway Protocol (IGP) Routes Over Traffic Engineering Tunnels", RFC 3906, DOI 10.17487/RFC3906, October 2004, <https://www.rfc-editor.org/info/rfc3906>.

[RFC5715] Shand, M. and S. Bryant, "A Framework for Loop-Free Convergence", RFC 5715, DOI 10.17487/RFC5715, January 2010, <https://www.rfc-editor.org/info/rfc5715>.

[RFC6976] Shand, M., Bryant, S., Previdi, S., Filsfils, C., Francois, P., and O. Bonaventure, "Framework for Loop-Free Convergence Using the Ordered Forwarding Information Base (oFIB) Approach", RFC 6976, DOI 10.17487/RFC6976, July 2013, <https://www.rfc-editor.org/info/rfc6976>.

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

Acknowledgements

We would like to thank the authors of [RFC6976] for introducing the concept of ordered convergence: Mike Shand, Stewart Bryant, Stefano Previdi, and Olivier Bonaventure.

Authors' Addresses

Stephane Litkowski
Orange

Email: stephane.litkowski@orange.com

Bruno Decraene
Orange

Email: bruno.decraene@orange.com

Clarence Filsfils
Cisco Systems

Email: cfilsfil@cisco.com

Pierre Francois
Individual Contributor

Email: pfrpfr@gmail.com