Network Working Group
Internet Draft
Intended status: Standards Track
Expires: August 15, 2025

W. Cheng
L. Gong
China Mobile
C. Lin
M. Chen
New H3C Technologies
February 15, 2025

OSPF Adjacency Suppression
draft-cheng-lsr-ospf-adjacency-suppress-04

Abstract

This document describes a mechanism for a router to instructs its neighbors to suppress advertising the adjacency to it until link-state database synchronization and LSA reoriginating are complete. This minimizes transient routing disruption when a router restarts from unplanned outages.

Status of this Memo

Copyright Notice

Table of Contents

1. Introduction

A router that is restarting from unplanned outages may not have maintained forwarding function state. Since this is not the first time the router has started, copies of LSAs generated by this router in its previous incarnation may exist in the link-state databases of other routers in the network. These copies are likely to appear "newer" than LSAs initially generated by the starting router due to the reinitialization of LSA sequence numbers by the starting router. So, without requesting the starting router to update its LSAs, the neighbors of the starting router may transition to "Full" state and route the traffic through the starting router. This may cause temporary blackholes to occur until the normal operation of the update process causes the starting router to reoriginate and flood copies of its own LSAs with higher sequence numbers.

This document describes OSPF extensions for adjacency suppression. This OSPF protocol extension provides functionality similar to the SA bit of Restart TLV in IS-IS [RFC8706]. With the proposed mechanism, the starting router’s neighbors will suppress advertising an adjacency to the starting router until the starting router has been able to propagate newer versions of LSAs, so that the temporary blackholes can be avoided.

1.1. 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.

2. Problem

2.1. Scenario of Two Router Network

Assume that in a simple OSPF network with two routers A-B, router A restarts from an unplanned outage, as shown in Figure 1.

An external route 10.1.1.0/24 is advertised by router A for some connected servers. After router A restarts, that external route is deleted and will not be advertised until data plane is ready to transfer packets for those servers. However, The old copies of LSAs generated by router A still exists in the link-state databases of router B, such as the Router-LSA with adjacency A->B and the External-LSA with 10.1.1.0. The restarting router A reinitializes LSA sequence numbers, hence the old copies appear to be "newer". Without requesting router A to update its LSAs, router B will transition to "Full" state and route the traffic through router A. A temporary blackhole occurs.

```
         External 10.1.1.0/24 (Down)
            |
            |
         Router A  ----------------------- Router B
       (Restarting)
            | --------- 1-way Hello --------> | Init
            |                                 |
    ExStart | <-------- 2-way Hello --------- |
            |                                 |
            | --------- 2-way Hello --------> | Exstart
            |                                 |
   ExChange | <------------ DD -------------> | ExChange
            |                                 |
    Loading | <-------- LS Request ---------> | Loading
            |                                 |
            | <-------- LS Update ----------> | Full ---
            |                                 |       ^
            | --------- LS Request ---------> |       |
            |  (Type 1, Seq M; Type 5, Seq N) |       |
            |                                 |  Blackhole
       Full | <-------- LS Update ----------- |  to 10.1.1.0
            |    (Type 1, Seq M, with A->B)   |       |
            |  (Type 5, Seq N, with 10.1.1.0) |       |
            |                                 |       v
            | --------- LS Update ----------> | --------
            |       (Self, Reorigniate)       |
              (Type 1, Seq M+1, without A->B)
                  (Type 5, Seq N, MaxAge)
```

Figure 1: Restarting Scenario of Two Router Network

The above procedure can also be summarized as the following steps:

- o Step 1.1: Router A restarts from unplanned outage and router B has the old LSA of router A in its link-state database;

- o Step 1.2: Router B reaches the Full state, and update its Router-LSA to advertise the adjacency B->A;

- o Step 1.3: Temporary blackhole occurs;

- o Step 1.4: Router B receives the reoriginated LSAs of router A;

- o Step 1.5: Temporary blackhole disappears.

Especially when router B has many more LSAs than router A, the time between Step 1.2 and Step 1.4 will be prolonged, and the impact of blackhole could be more significant.

In addition to external routes, other types of routes which have old copies on neighbor may have the same problem during restarting.

2.2. Scenario of Network with More Router

Assume that there are more routers in the network, as shown in the following figure. Router C represents the rest of the network attached to router B.

```
     +-----+        +-----+        +-----+
     |Rtr A|--------|Rtr B|--------|Rtr C|
     +-----+        +-----+        +-----+
  (Restarting)             (Rest of the Network)
```

Figure 2: Restarting Scenario of Network with More Routers

From the perspective of router C, a temporary blackhole may also occur when the following order comes:

o Step 2.1: Router A restarts from unplanned outage and router C has the old LSA of router A in its link-state database;

o Step 2.2: Router C receives the new Router-LSA of B advertising the adjacency B->A;

o Step 2.3: Temporary blackhole occurs;

o Step 2.4: Router C receives the reoriginated LSAs of router A without the adjacency A->B;

o Step 2.5: Temporary blackhole disappears.

The above procedure is likely to occur under certain conditions, such as packet loss, out of order, MinLSInterval, or MinLSArrival, because the sequence of the flooding process cannot be controlled precisely.

2.3. Scenario of Router Starting

```
     +-----+        +-----+        +-----+
     |Rtr A|--------|Rtr B|--------|Rtr C|
     +-----+        +-----+        +-----+
   (Starting)              (Rest of the Network)
```

Figure 3: Starting Scenario of Network with More Routers

When Router A initially starts up, its internal forwarding processes may not yet be ready or its dependencies may not have recovered, and therefore it cannot forward traffic properly. As a result, Router A prefers that other routers do not forward any traffic to it until all its internal processes have been completed. Existing mechanisms, such as stub-router and reverse metric, cannot prevent remote routers from forwarding traffic to this starting device.

3. Solution

The solution proposed in this document is to allow the restarting router to control the timing for its neighbor to advertise adjacency after FULL state.

- o Step 3.1: The restarting router signals suppressing adjacency to its neighbor;
- o Step 3.2: The neighboring router suppresses the advertisement of the adjacency to the starting router (even if it transitions to the FULL state during this period);
- o Step 3.3: The restarting router reoriginates and floods its own LSAs;
- o Step 3.4: The restarting router stops signaling suppressing adjacency to its neighbor;
- o Step 3.5: The neighboring router advertises the adjacency to the restarting router.

The proposed solution is similar with the mechanism of the SA bit of Restart TLV in IS-IS [RFC8706]. The OSPF signaling of suppressing adjacency is called the SA-Indicator, which will be specified in Section 4.

3.1. Sending the SA-Indicator

When a router is starting, it starts a timer T-SA and sends the SA-Indicator to its neighbors. The timer T-SA is canceled after the following conditions are met:

- o Synchronization of the link-state database is complete.
- o Reoriginating its own LSAs is complete (with additional delay).
- o The local forwarding plane is ready.
- o External dependencies are resolved, such as waiting for BGP convergence.

When the timer T-SA has expired or been canceled, the starting router MUST clear the SA-Indicator.

3.2. Receiving the SA-Indicator

When a router receives an SA-Indicator, if there exists on this interface an adjacency in the FULL state with the same Router ID, then the router MUST suppress advertisement of the adjacency to the neighbor in its own LSAs. In the case of broadcast and NBMA links, the Designated Routers are responsible for the suppressing of adjacency advertisement.

Until the SA-Indicator is cleared, the adjacency advertisement MUST continue to be suppressed. During that period, if the neighbor transitions to the FULL state, the new adjacency MUST NOT be advertised.

Besides, a router that suppresses advertisement of an adjacency MUST NOT use this adjacency when performing its SPF calculation.

4. OSPF Extensions for SA-Indicator

This Section defines the extensions for OSPF protocol to advertise SA-Indicator (Suppressing Adjacency Indicator). The advertising of SA-Indicator has several options. Section 4.1 describes how to advertise SA-Indicator in Hello packets with LLS [RFC5613]. Section 4.2 describes how to advertise SA-Indicator in link-local Opaque-LSAs [RFC5250].

4.1. Advertising SA-Indicator in Hello Packets with LLS

There are two possible positions in the OSPF LLS [RFC5613] to carry the SA-Indicator, which are specified in Section 4.1.1 and 4.1.2.

4.1.1. Option A: Extended Options and Flags TLV

The SA-Indicator can be carried in the Type 1 Extended Options and Flags TLV [RFC5613] as a new SA-bit. This bit is defined for the LLS block included in Hello packets and instructs the receiver to suppress advertising an adjacency to the sender.

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|             1                 |            4                  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                  Extended Options and Flags                   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+

Extended Options Bit:
  0x00000001: LR-bit
  0x00000002: RS-bit
  0x00000004: I-bit
  0x00000008: F-bit
  0x00000010: B-bit
  0x00000020: FR-bit
  TBD       : SA-bit
```

Figure 4: Format of the Extended Options and Flags TLV

4.1.2. Option B: Reverse Metric TLV

The SA-Indicator can be carried in the Type 19 Reverse Metric TLV [RFC9339] as a new SA-bit. This bit is defined for the LLS block included in Hello packets and instructs the receiver to suppress advertising an adjacency to the sender.

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|             19                |               4               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     MTID      |     Flags     |        Reverse Metric         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+

Flags:
  0x00000001: H-bit
  0x00000002: O-bit
  TBD       : SA-bit
```

Figure 5: Format of the Reverse Metric TLV

4.1.3. Operations

- o Set the SA-Indicator: Send Hello packets containing the LLS block with the Extended Options and Flags TLV or Reverse Metric TLV that has the SA-bit set.

- o Unset the SA-Indicator: Send Hello packets with the SA-bit clear.

4.2. Advertising SA-Indicator in Link-local Opaque-LSAs

The restarting router can originate link-local Opaque-LSAs, called the SA-LSAs as defined in Section 4.2.1, to advertise the SA-Indicator to its neighbors.

4.2.1. Option C: SA-LSA

The SA-LSA is a link-local scoped Opaque-LSA. The Opaque Type is TBA and the Opaque ID equal to 0. SA-LSAs are originated by a router that wishes the receiver to suppress advertising an adjacency to the originator.

Each SA-LSA has an LS age field set to 0 when the LSA is first originated; the current value of the LS age then indicates how long ago the restarting router made its request for suppressing adjacency. The body of the LSA is TLV-encoded.

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|            LS age             |     Options   |       9       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|      TBA      |                    0                          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Advertising Router                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     LS sequence number                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|         LS checksum           |             length            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                                               |
+-                            TLVs                             -+
|                             ...                               |
```

Figure 6: Format of the SA-LSA

The format of the TLVs within the body of a SA-LSA is as following:

```
  0                   1                   2                   3
  0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |              Type             |             Length            |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |                            Value...                           |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

The following is the list of TLVs that can appear in the body of a SA-LSA:

- o IP interface address (Type=1, length=4). The router’s IP interface address on the subnet associated with the SA-LSA. Required on broadcast, NBMA and Point-to-MultiPoint segments, where the neighbor uses the IP interface address to identify the restarting router.

DoNotAge is never set in a SA-LSA.

SA-LSAs have link-local scope because they only need to be seen by the router’s direct neighbors.

4.2.2. Operations

- o Set the SA-Indicator: Originate the SA-LSA.

- o Unset the SA-Indicator: Flush the SA-LSA.

5. Backward Compatibility

The described technique requires cooperation from neighboring routers. If a router does not support this technique, it will ignore the SA-Indicator and advertise the adjacency when the neighbor transitions to the FULL state. As a result, the behavior would be the same as without this specification.

6. Management Considerations

Support of the suppressing adjacency SHOULD be based on local configuration, and the interval of the timer T-SA SHOULD be configurable.

7. Security Considerations

TBD.

8. IANA Considerations

TBD.

9. References

9.1. Normative References

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

[RFC5250] Berger, L., Bryskin, I., Zinin, A., and R. Coltun, "The OSPF Opaque LSA Option", RFC 5250, DOI 10.17487/RFC5250, July 2008, <https://www.rfc-editor.org/info/rfc5250>.

[RFC5613] Zinin, A., Roy, A., Nguyen, L., Friedman, B., and D. Yeung, "OSPF Link-Local Signaling", RFC 5613, DOI 10.17487/RFC5613, August 2009, <https://www.rfc-editor.org/info/rfc5613>.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, May 2017

[RFC9339] Talaulikar, K., Psenak, P., and H. Johnston, " OSPF Reverse Metric", RFC 9339, DOI 10.17487/RFC9339, December 2022, <https://www.rfc-editor.org/info/rfc9339>.

9.2. Informative References

[RFC8706] Ginsberg, L., and P. Wells, "Restart Signaling for IS-IS", RFC 8706, DOI 10.17487/RFC8706, February 2020, <https://www.rfc-editor.org/info/rfc8706>.

Acknowlegements

The authors would like to acknowledge Les Ginsberg for highlighting the problems on remote routers.

Authors’ Addresses

Weiqiang Cheng
China Mobile
China
Email: chengweiqiang@chinamobile.com

Liyan Gong
China Mobile
China
Email: gongliyan@chinamobile.com

Changwang Lin
New H3C Technologies
China
Email: linchangwang.04414@h3c.com

Mengxiao Chen
New H3C Technologies
China
Email: chen.mengxiao@h3c.com