Workgroup: Network Working Group
Internet-Draft:
draft-hegde-lsr-ospf-better-idbx-01
Published: 16 October 2023
Intended Status: Informational
Expires: 18 April 2024
Authors: S. Hegde   A. Przygienda   A. Lindem
Juniper   Juniper   LabN Consulting LLC

Improved OSPF Database Exchange Procedure

Abstract

When an OSPF router undergoes restart, previous instances of LSAs belonging to that router may remain in the databases of other routers in the OSPF domain until such LSAs are aged out. Hence, when the restarting router joins the network again, neighboring routers re-establish adjacencies while the restarting router is still bringing-up its interfaces and adjacencies and generates LSAs with sequence numbers that may be lower than the stale LSAs. Such stale LSAs may be interpreted as bi-directional connectivity before the initial database exchanges are finished and genuine bi-directional LSA connectivity exists. Such incorrect interpretation may lead to, among other things, transient traffic packet drops. This document suggests improvements in the OSPF database exchange process to prevent such problems due to stale LSA utilization. The solution does not preclude changes in the existing standard but presents an extension that will prevent this scenario.

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.

Status of This Memo

at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on 18 April 2024.

Copyright Notice

Table of Contents

1. Introduction

When an OSPF [RFC2328] router restarts, its stale LSAs are left in the database of other routers in the OSPF domain until the LSAs are aged out either intentionally or by the LSA age elapsing. The stale Router LSA can contain links to all the neighbors that had Full adjacencies before the router restarted.

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

Figure 1: OSPF Network

Figure 1 shows a very simple OSPF network. In case of C undergoing restart that does not generate purges, the other routers in the domain will hold the stale LSA of Router C in their database. The stale LSA may have links to B and E, which represents the topology of C before it went down. When C restarts again, it initiates the database exchange process with B and E. B and E may have C's stale LSA with a higher sequence number in their database than the new ones originated by C and hence assume this is latest copy, successively bringing up the adjacency with C, and transitioning to Full state. Based on C's Stale LSA having LSA links to B and E, the Shortest Path First (SPF) back-link check is satisfied and B and E update their routing table to point to C. This may cause C to drop this traffic as C may not have all its previous adjacencies up and all LSAs in place to correctly compute the necessary routes.

2. Solution

The database exchange procedure from [RFC2328] section 7.2 is extended with additional constraints to prevent an OSPF router from transitioning to Full state when it has stale LSAs originated by the database exchange neighbor in its Link State Database (LSDB).

During Database exchange, when a router receives an LSA from the neighbor for which such neighbor is the originator of the LSA and the sequence number of the LSA is smaller than the sequence number of its own database copy, the receiving router marks its database copy as stale. This document proposes to create a new LSA list called "Stale LSA List". This list is created on a per neighbor basis and resembles the "LS Request List", the difference being LS Request messages are not sent for stale LSAs. The router creates a "Stale LSA List" for this neighbor and adds stale LSA to the Stale LSA List for the neighbor. This LSA MUST NOT be removed from the Stale LSA list and the adjacency FSM MUST NOT transition to Full state until an LSA with a sequence number greater than its own database copy is received (or strictly speaking, a "more recent" LSA). The Stale LSA List cleanup procedures follow the LSRequest list cleanup procedures as described in [RFC2328]

2.1. Example

Figure 2 provides an example of C restarting having originated an LSA with sequence number Y before. After restarting C originates the same LSA with sequence number X where X < Y since it is not aware of existence of version X yet.

```
       C-----------------E

   DBD: LSA A origin:C  ------->
       Sequence:X
                           DBD: LSA A Origin:C
                   <------       Sequence:Y
   LSReq LSA A,
   origin C       ----->
                         Modified Procedure
                  <------ Add LSA A to Stale LSA List

                  <------- LSA  A, Origin C, Seq:Y
   C re-originating self LSA
   LSA A, Origin C, -------> Remove Stale
   Seq:Y+1                    LSA A, Origin C
                              From Stale LSA List,

                              Bring adjacency to Full state if
                              both LS Request list and Stale LSA
                              list are empty.
```

Figure 2: Modified Database Exchange Procedure

As shown in figure Figure 2 above, E originates LSReq with Sequence number X but waits until the LSA with sequence number Y+1 (or strictly speaking, an LSA that compares as newer to the one it holds) arrives. As the LSA is still in the Stale LSA List, the adjacency will remain in Loading state and will not move to Full state. All the neighbors of the restarting routers hold the neighbor FSM in Loading state and do not transition to Full state until the stale LSA is replaced with the new LSA with higher sequence number than the stale LSA. This ensures that other routers in the network do not compute a path through the restarting router since they cannot satisfy the bi-directionality condition in SPF computations.

3. IANA Considerations

No IANA Considerations

4. Security Considerations

5. Contributors

6. References

6.1. Informative References

6.2. Normative References

[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>.

Authors' Addresses

Shraddha Hegde
Juniper
India

Email: shraddha@juniper.net

Tony Przygienda
Juniper
1133 Innovation Way
Sunnyvale, CA
United States of America

Email: prz@juniper.net

Acee Lindem
LabN Consulting LLC
301 Midenhall Way
Cary, North Carolina
United States of America

Email: acee.ietf@gmail.com