LDP IGP Synchronization for Broadcast Networks

Abstract

RFC 5443 describes a mechanism to achieve LDP IGP synchronization to
prevent black-holing traffic (e.g., VPN) when an Interior Gateway
Protocol (IGP) is operational on a link but Label Distribution
Protocol (LDP) is not. If this mechanism is applied to broadcast
links that have more than one LDP peer, the metric increase procedure
can only be applied to the link as a whole but not to an individual
peer. When a new LDP peer comes up on a broadcast network, this can
result in loss of traffic through other established peers on that
network. This document describes a mechanism to address that use-
case without dropping traffic. The mechanism does not introduce any
protocol message changes.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

   In RFC 5443 [LDP-IGP-SYNC], when [LDP] is not fully operational on a
   link, the IGP advertises the link with maximum cost to avoid any
   transit traffic on the link if possible. When LDP becomes
   operational, i.e., all the label bindings have been exchanged, the
   link is advertised with its correct cost. This tries to ensure that
   the LDP Label Switch Path (LSP) is available all along the IGP
   shortest path. The mechanisms in [LDP-IGP-SYNC] have limitations
   when applied to a broadcast link. These are described in Section 3.
   A solution is defined in Section 4.

2. Conventions Used in This Document

   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].

3. Problem Statement

   On broadcast networks, a router's Link State Advertisement (LSA)
   contains a single cost to the broadcast network rather than a
   separate cost to each peer on the broadcast network. The operation
   of the mechanism in [LDP-IGP-SYNC] is analyzed using the sample
   topology in Figure 1, where routers A, B, C, and E are attached to a

common broadcast network. Say all links in that topology have a cost
of 1 except the link A-PE3, which has a cost of 10. The use-case
when router B's link to the broadcast network comes up is analyzed.
Before that link comes up, traffic between PE1 and PE2 flows along
the bi-directional path PE1-A-C-D-PE2, and traffic between PE1 and
PE3 flows along the bi-directional path PE1-A-E-PE3.

```
                                  |   +---+             +---+
                             ----| B |-----------|PE2|
                                  |   +---+             +---+
       +---+      +---+           |
       |PE1|----| A |----         |
       +---+      +---+           |
                     |            |   +---+     +---+       |
                     |       ----| C |----| D |----+
                     |            |   +---+     +---+
                     |            |
                     |            |
                     |            |   +---+
                     |       ----| E |-------------+
                     |            |   +---+             |
                     |            |                     |
                     |                                  |
                     |                                +---+
                     +--------------------------|PE3|
                                                      +---+
```
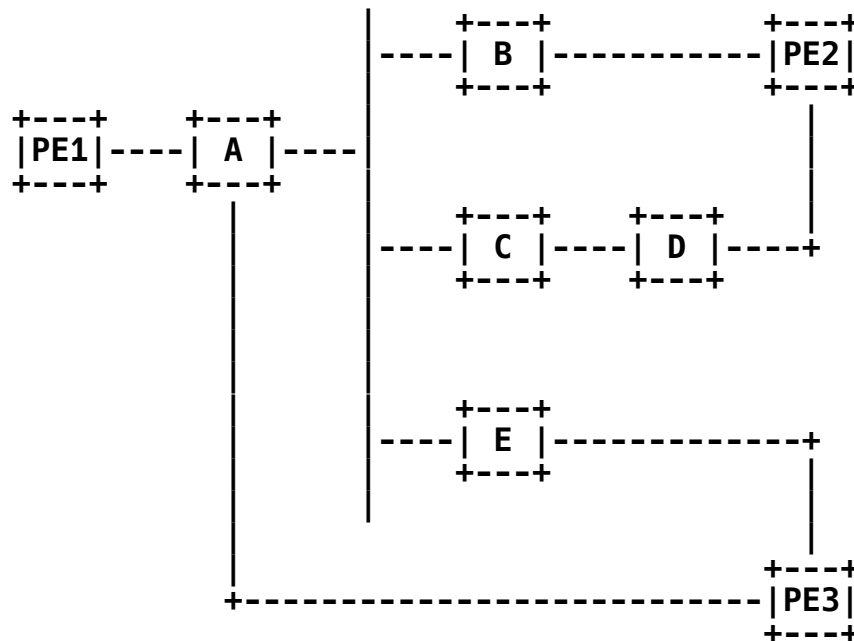
                Figure 1: LDP IGP Sync on a Broadcast Network

In one interpretation of the applicability of [LDP-IGP-SYNC] to
broadcast networks, when a new router is discovered on a broadcast
network, that network should avoid transit traffic until LDP becomes
operational between all routers on that network. This can be
achieved by having all the attached routers advertise maximum cost to
that network. This should result in traffic that is being sent via
that broadcast network to be diverted. However, traffic might be
inadvertently diverted to the link that just came up. Until LDP
becomes operational, that traffic will be black-holed. An additional
problem is route churn in the entire network that results in traffic
that should be unaffected taking sub-optimal paths until the high-
cost metric is reverted to the normal cost. In Figure 1, when B's
link to the broadcast network comes up and it is discovered by
routers A, C and E, then A, B, C, and E can all start advertising
maximum cost to the broadcast network. A will have B as next-hop to
PE2 and will not have a LDP LSP to PE2, resulting in VPN traffic from
PE1 to PE2 to be black-holed at A. The route churn at A also results
in traffic between PE1 and PE3 to be unnecessarily diverted to the

   sub-optimal path PE1-A-PE3 until the maximum-cost advertisement is
   reverted to the normal cost.

   This interpretation has the additional complexity of requiring the
   maximum-cost advertisement to be reverted by all routers after LDP
   peering between all the routers on the broadcast network is
   operational. This is non-trivial and needs coordination between all
   the routers.

   In another alternative interpretation of the applicability of
   [LDP-IGP-SYNC] to broadcast networks, only the router whose link to
   the broadcast network comes up advertises maximum cost for that link,
   but other routers continue to advertise the normal cost. In Figure
   1, when B's link to the broadcast network comes up, it advertises a
   high cost to the broadcast network. After the IGP has converged but
   the LDP peering A-B is not yet operational, A will have B as the
   next-hop for PE2 and will not have a LDP LSP to PE2. Since A's cost
   to reach B is not high, A-B-PE2 becomes the shortest path. VPN
   traffic from PE1 to PE2 will be dropped at A.

4. Solution

   The problem described above exists because the Link State Database
   (LSDB) of the IGP does not describe a link coming up on a broadcast
   network with a high bi-directional cost to all other routers on that
   broadcast network. A broadcast network is advertised as a pseudonode
   containing a list of routers to which the broadcast network is
   connected, and the cost of all these links from the pseudonode to
   each router is zero when computing SPF (Shortest Path First).

   The solution proposed below removes the link that is coming up from
   the LSDB unless absolutely necessary. Only the router whose link is
   coming up plays a role in ensuring this. The other routers on the
   broadcast network are not involved. The following text describes
   this in more detail.

   During the intra-area SPF algorithm execution, an additional
   computation is made to detect an alternate path to a directly
   connected network that does not have any IGP adjacencies.

   If a router has a directly connected network that does not have an
   alternate path to reach it, then the interface to that network is a
   "cut-edge" in the topology for that router. When a "cut-edge" goes
   down, the network is partitioned into two disjoint sub-graphs. This
   property of whether or not an interface is a "cut-edge" is used when
   an IGP adjacency comes up on that interface. The method to determine
   whether an interface is a "cut-edge" is described in Appendix A.

During IGP procedures, when the router's first adjacency to the
broadcast network is coming up and the LSA is about to be updated
with a link to the pseudonode of the broadcast interface, a check is
made whether that interface is a "cut-edge". If it is not a
"cut-edge", then the updating of the LSA with that link to the
pseudonode is postponed until LDP is operational with all the LDP
peers on that broadcast interface. After LDP is operational, the LSA
is updated with that link to the pseudonode (and the LSA is flooded).
If the interface is a "cut-edge", then the updating of the LSA MUST
NOT be delayed by LDP's operational state. Note that the IGP and LDP
adjacency bring-up procedures are unchanged. The conditional check
of whether the interface is a "cut-edge" must be done just before the
adjacency is about to be reflected in the LSA.

If the IGP is [OSPF], the Router-LSA is not updated with a "Link Type
2" (link to transit network) for that subnet until LDP is operational
with all neighboring routers on that subnet.

Similarly, if the IGP is [IS-IS], the "Link State PDU" is updated
with an "IS Reachability TLV" (or an "Extended IS Reachability TLV")
to the pseudonode after LDP is operational with all neighboring
routers on that subnet.

Note that this solution can be introduced in a gradual manner in a
network without any backward compatibility issues.

5. Scope

This document is agnostic to the method that detects LDP to be
operational with a neighbor. It does not define any new method to
detect that LDP is operational. At the time of publishing this
document, LDP End-of-LIB [LDP-EOL] seems to be the preferred method.

Issues arising out of LDP not being configured on some routers or on
some interfaces are not specific to the method described in this
document and are considered outside the scope of this solution.

6. Applicability

The method described in this document can be easily extended to
point-to-point (P2P) links. However, an implementation may continue
to apply the method described in [LDP-IGP-SYNC] to P2P links but
apply the method described in this document to broadcast networks.
Both methods can coexist in a network.

The techniques used in this document's solution enable LDP IGP
synchronization in many scenarios where one end of the IGP adjacency
does not support any LDP IGP sync method. This is an optional
benefit and is for further study. Some ways to apply this technique
to achieve that benefit are discussed in Appendix B.

7. Security Considerations

This document does not introduce any new security considerations
beyond those already described in [LDP-IGP-SYNC].

Note that in [LDP-IGP-SYNC], when a link is advertised with a high
metric, an alternate path with a large number of hops can result in
the end-to-end path having more than 255 hops and thus result in
unreachability. This fact could be exploited if control of metrics
falls into the hands of an attacker.

This problem can even exist in a plain IP network with a link-state
IGP. If the directly connected path has a higher metric than an
alternate path with Time to Live (TTL) greater than 255 hops, then
the standard SPF algorithm will conclude that the shortest path is
the alternate path although the neighboring node is unreachable
through this path. In this case, the link is advertised with its
normal metric yet there is unreachability in the network. Thus, this
document does not introduce any new issues beyond those in a standard
IGP-based IP network, and operators need to apply policy and security
to the techniques used to determine and distribute the metrics used
on links in their networks.

8. Conclusions

This document complements [LDP-IGP-SYNC] by providing a solution to
achieve LDP IGP synchronization for broadcast networks. It can also
coexist with that solution in a network that has a combination of P2P
links and broadcast networks. It can also be introduced into a
network without backward compatibility issues. The solution in this
document can also be used exclusively to achieve LDP IGP
synchronization since this solution applies to both P2P links and
broadcast networks.

This solution also has useful properties that can be optionally used
to achieve LDP IGP synchronization when only one end of the IGP
adjacency supports this solution but the other end supports neither
this solution nor the one in [LDP-IGP-SYNC].

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.

   [LDP-IGP-SYNC]
               Jork, M., Atlas, A., and L. Fang, "LDP IGP
               Synchronization", RFC 5443, March 2009.

   [LDP]   Andersson, L., Ed., Minei, I., Ed., and B. Thomas, Ed.,
               "LDP Specification", RFC 5036, October 2007.

   [OSPF]   Moy, J., "OSPF Version 2", STD 54, RFC 2328, April 1998.

   [IS-IS]   International Organization for Standardization,
               "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
               Standard 10589, 2002.

9.2. Informative References

   [LDP-EOL]   Asati, R., Mohapatra, P., Chen, E., and B. Thomas,
               "Signaling LDP Label Advertisement Completion", RFC 5919,
               August 2010.

Acknowledgments

Appendix A. Computation of "Cut-Edge"

   A "cut-edge" can be computed during an intra-area SPF run or by using
   results of the previous SPF run. If an SPF run was scheduled but is
   pending execution, that SPF MUST be executed immediately before any
   procedure checks whether an interface is a "cut-edge".

   An interface is considered a "cut-edge" if, during intra-area SPF
   (using Dijkstra's algorithm described in Section 16.1 of [OSPF]),
   there is no alternate path for the directly connected network.
   Alternately, a "cut-edge" can be detected by the last run of SPF if
   there is a lack of connectivity to the router-id of a directly
   connected peer via an alternate path. The router-id can be known
   during the adjacency bring-up process.

   A "cut-edge" computation should not require any extra SPF runs. It
   should not increase the algorithmic complexity of SPF.

Appendix B. Sync without Support at One End

   A useful property of the solution described in this document is that
   LDP IGP synchronization is achievable in many scenarios where one end
   of the IGP adjacency does not support any LDP IGP sync method.

   For P2P links (or broadcast links on which the IGP operates in P2P
   mode) the applicability is straightforward. An IGP can establish a
   P2P adjacency on a P2P link or a broadcast link with the IGP in P2P
   mode. When a P2P adjacency comes up, the end of the adjacency that
   supports the solution in this document would not advertise the link
   to the other router in its LSA unless the edge is a "cut-edge" or
   until LDP becomes operational. Hence, neither of the two routers
   will have IGP next-hop as the other router unless the link is a
   "cut-edge". Consider Figure 1 modified such that the broadcast
   network is replaced by P2P links between each of A, B, C, and E. Say
   link A-B is coming up, but only A has implemented the solution in
   this document whereas B has implemented neither the solution in this
   document nor the solution in [LDP-IGP-SYNC]. Since A's LSA does not
   advertise a link to B until LDP is operational, B does not have A as
   next-hop. After LDP is operational, A advertises the link to B in
   its LSA. Hence, there is no traffic loss due to LDP LSP not being
   present.

   For broadcast networks, the applicability is not straightforward and
   should be considered a topic for future study. One way is for the
   designated router (DR) to stop advertising the link in the pseudonode
   to the router whose link is coming up until LDP is operational.

Authors' Addresses

    Sriganesh Kini (editor)
    Ericsson
    300 Holger Way
    San Jose, CA 95134
    EMail: sriganesh.kini@ericsson.com

    Wenhu Lu (editor)
    Ericsson
    300 Holger Way
    San Jose, CA 95134
    EMail: wenhu.lu@ericsson.com