Network Working Group
Request for Comments: 5443
Category: Informational

M. Jork
GENBAND
A. Atlas
British Telecom
L. Fang
Cisco Systems, Inc.

March 2009

LDP IGP Synchronization

Status of This Memo

This memo provides information for the Internet community. It does not specify an Internet standard of any kind. Distribution of this memo is unlimited.

Copyright Notice

Abstract

In certain networks, there is dependency on the edge-to-edge Label Switched Paths (LSPs) setup by the Label Distribution Protocol (LDP), e.g., networks that are used for Multiprotocol Label Switching (MPLS) Virtual Private Network (VPN) applications. For such applications, it is not possible to rely on Internet Protocol (IP) forwarding if the MPLS LSP is not operating appropriately. Blackholing of labeled traffic can occur in situations where the Interior Gateway Protocol (IGP) is operational on a link on which LDP is not. While the link could still be used for IP forwarding, it is not useful for MPLS forwarding, for example, MPLS VPN applications or Border Gateway Protocol (BGP) route-free cores. This document describes a mechanism to avoid traffic loss due to this condition without introducing any protocol changes.

Table of Contents

1. Introduction

LDP [RFC5036] establishes MPLS LSPs along the shortest path to a destination as determined by IP forwarding. In a common network design, LDP is used to provide Label Switched Paths throughout the complete network domain covered by an IGP such as Open Shortest Path First (OSPF) [RFC2328] or Intermediate System to Intermediate System (IS-IS) [ISO.10589.1992]; i.e., all links in the domain have IGP as well as LDP adjacencies.

A variety of services a network provider may want to deploy over an LDP-enabled network depend on the availability of edge-to-edge label switched paths. In a layer 2 (L2) or layer 3 (L3) VPN scenario, for example, a given Provider-Edge (PE) router relies on the availability of a complete MPLS forwarding path to the other PE routers for the VPNs it serves. This means that all the links along the IP shortest path from one PE router to the other need to have operational LDP sessions, and the necessary label binding must have been exchanged over those sessions. If only one link along the IP shortest path is

not covered by an LDP session, a blackhole exists and services depending on MPLS forwarding will fail. This might be a transient or a persistent error condition. Some of the reasons for this could be:

- A configuration error.

- An implementation bug.

- The link has just come up and has an IGP adjacency but LDP has either not yet established an adjacency or session, or has not yet distributed all the label bindings.

The LDP protocol has currently no way to correct the issue. LDP is not a routing protocol; it cannot re-direct traffic to an alternate IGP path.

1.1. 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].

2. Proposed Solution

The problem described above exists because LDP is tied to IP-forwarding decisions but no coupling between the IGP and LDP operational state on a given link exists. If IGP is operational on a link but LDP is not, a potential network problem exists. So the solution described by this document is to discourage a link from being used for IP forwarding as long as LDP is not fully operational.

This has some similarity to the mechanism specified in [RFC3137], which allows an OSPF router to advertise that it should not be used as a transit router. One difference is that [RFC3137] raises the link costs on all (stub) router links, while the mechanism described here applies on a per-link basis.

In detail: when LDP is not "fully operational" (see below) on a given link, the IGP will advertise the link with maximum cost to avoid any transit traffic over it. In the case of OSPF, this cost is LSInfinity (16-bit value 0xFFFF), as proposed in [RFC3137]. In the case of ISIS, the maximum metric value is 2^24-2 (0xFFFFFE). Indeed, if a link is configured with 2^24-1 (the maximum link metric per [RFC5305]), then this link is not advertised in the topology. It is important to keep the link in the topology to allow IP traffic to use the link as a last resort in case of massive failure.

LDP is considered fully operational on a link when an LDP hello adjacency exists on it, a suitable associated LDP session (matching the LDP Identifier of the hello adjacency) is established to the peer at the other end of the link, and all label bindings have been exchanged over the session. At the present time, the latter condition cannot generally be verified by a router and some estimation may have to be used. A simple implementation strategy is to use a configurable hold-down timer to allow LDP session establishment before declaring LDP fully operational. The default timer is not defined in this document due to concerns of the large variations of the Label Information Base (LIB) table size and equipment capabilities. In addition, there is a current work in progress on LDP End-of-LIB as specified in [End-of-LIB], which enables the LDP speaker to signal the completion of its initial advertisement following session establishment. When LDP End-of-LIB is implemented, the configurable hold-down timer is no longer needed. The neighbor LDP session is considered fully operational when the End-of-LIB notification message is received.

This is typically sufficient to deal with the link when it is being brought up. LDP protocol extensions to indicate the complete transmission of all currently available label bindings after a session has come up are conceivable, but not addressed in this document.

The mechanism described in this document does not entail any protocol changes and is a local implementation issue.

The problem space and solution specified in this document have also been discussed in an IEEE Communications Magazine paper [LDP-Fail-Rec].

3. Applicability

In general, the proposed procedure is applicable in networks where the availability of LDP-signaled MPLS LSPs and avoidance of blackholes for MPLS traffic are more important than always choosing an optimal path for IP-forwarded traffic. Note however that non-optimal IP forwarding only occurs for a short time after a link comes up or when there is a genuine problem on a link. In the latter case, an implementation should issue network management alerts to report the error condition and enable the operator to address it.

Example network scenarios that benefit from the mechanism described here are MPLS VPNs and BGP-free core network designs where traffic can only be forwarded through the core when LDP forwarding state is available throughout.

The usefulness of this mechanism also depends on the availability of alternate paths with sufficient bandwidth in the network should one link be assigned to the maximum cost due to the unavailability of LDP service over it.

On broadcast links with more than one IGP/LDP peer, the cost-out procedure can only be applied to the link as a whole and not to an individual peer. So a policy decision has to be made whether the unavailability of LDP service to one peer should result in the traffic being diverted away from all the peers on the link.

4. Interaction with TE Tunnels

In some networks, LDP is used in conjunction with RSVP-TE, which sets up traffic-engineered tunnels. The path computation for the TE tunnels is based on the TE link cost that is flooded by the IGP in addition to the regular IP link cost. The mechanism described in this document should only be applied to the IP link cost to prevent unnecessary TE tunnel reroutes.

In order to establish LDP LSPs across a TE tunnel, a targeted LDP session between the tunnel endpoints needs to exist. This presents a problem very similar to the case of a regular LDP session over a link (the case discussed so far): when the TE tunnel is used for IP forwarding, the targeted LDP session needs to be operational to avoid LDP connectivity problems. Again, raising the IP cost of the tunnel while there is no operational LDP session will solve the problem. When there is no IGP adjacency over the tunnel and the tunnel is not advertised as a link into the IGP, this becomes a local issue of the tunnel headend router.

5. Security Considerations

A Denial-of-Service (DoS) attack that brings down LDP service on a link or prevents it from becoming operational on a link could be one possible cause of LDP-related traffic blackholing. This document does not address how to prevent LDP session failure. The mechanism described here prevents the use of the link where LDP is not operational while IGP is. Assigning the IGP cost to maximum on such a link should not introduce new security threats. The operation is internal to the router to allow LDP and IGP to communicate and react. Making many LDP links unavailable, however, is a security threat that can cause dropped traffic due to limited available network capacity. This may be triggered by operational error or implementation error. These errors are considered general security issues and implementors should follow the current best security practice [MPLS-GMPLS-Sec].

6. References

6.1. Normative References

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

[RFC2328] Moy, J., "OSPF Version 2", STD 54, RFC 2328, April 1998.

[RFC5036] Andersson, L., Ed., Minei, I., Ed., and B. Thomas, Ed., "LDP Specification", RFC 5036, October 2007.

6.2. Informative References

[End-of-LIB] Asati, R., LDP End-of-LIB, Work in Progress, January 2009.

[ISO.10589.1992] International Organization for Standardization, "Intermediate system to intermediate system intra-domain-routing routine information exchange protocol for use in conjunction with the protocol for providing the connectionless-mode Network Service (ISO 8473)", ISO Standard 10589, 1992.

[LDP-Fail-Rec] Fang, L., Atlas, A., Chiussi, F., Kompella, K., and G. Swallow, "LDP Failure Detection and Recovery", IEEE Communications Magazine, Vol.42, No.10, October 2004.

[MPLS-GMPLS-Sec] Fang. L., Ed., "Security Framework for MPLS and GMPLS Networks", Work in Progress, November 2008.

[RFC3137] Retana, A., Nguyen, L., White, R., Zinin, A., and D. McPherson, "OSPF Stub Router Advertisement", RFC 3137, June 2001.

[RFC5305] Li, T. and H. Smit, "IS-IS Extensions for Traffic Engineering", RFC 5305, October 2008.

7. Acknowledgments

The authors would like to thank Bruno Decraene for his in-depth discussion and comments, Dave Ward for his helpful review and input, and Loa Andersson, Ross Callon, Amanda Baber, Francis Dupont, Donald Eastlake, Russ Housley, Pasi Eronen, Dan Romascanu, Bin Mo, Lan Zheng, Bob Thomas, and Dave Ojemann for their reviews and comments.

Authors' Addresses

Markus Jork
GENBAND
3 Federal St.
Billerica, MA 01821
USA

EMail: Markus.Jork@genband.com

Alia Atlas
British Telecom

EMail: alia.atlas@bt.com

Luyuan Fang
Cisco Systems, Inc.
300 Beaver Brook Road
Boxborough, MA 01719
USA

EMail: lufang@cisco.com
Phone: 1 (978) 936-1633