

Network Working Group
Internet Draft
Category: Informational
Expires: May 2008

M. Jork
NextPoint Networks
Alia Atlas
British Telecom
L. Fang
Cisco Systems, Inc.

November 3, 2008

LDP IGP Synchronization
draft-ietf-mpls-ldp-igp-sync-03.txt

Status of this Memo

By submitting this Internet-Draft, each author represents that any applicable patent or other IPR claims of which he or she is aware have been or will be disclosed, and any of which he or she becomes aware will be disclosed, in accordance with [Section 6 of BCP 79](#).

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

The list of current Internet-Drafts can be accessed at
<http://www.ietf.org/ietf/lid-abstracts.txt>

The list of Internet-Draft Shadow Directories can be accessed at
<http://www.ietf.org/shadow.html>.

Copyright Notice

Copyright (C) The IETF Trust (2008).

Abstract

In certain networks there is a dependency on edge-to-edge LSPs setup by LDP, e.g. networks that are used for MPLS VPN applications. For such applications it is not possible to rely on IP forwarding if the MPLS LSP is not operating appropriately. Blackholing of labeled traffic can occur in situations where the IGP is operational on a link but LDP is not operational on that link. While the link could still be used for IP forwarding, it is not useful for MPLS

LDP IGP Synchronization

November 2008

forwarding, for example, MPLS VPN; BGP route free core; or IP address carried in the packet is out of the [RFC1918](#) space. This document describes a mechanism to avoid traffic loss due to this condition without introducing any protocol changes.

Table of Contents

1.	Introduction.....	2
2.	Proposed Solution.....	3
3.	Applicability.....	4
4.	Interaction With TE Tunnels.....	5
5.	Security Considerations.....	5
6.	IANA Considerations.....	5
7.	Normative References.....	6
8.	Informational References.....	6
9.	Author's Addresses.....	6
10.	Acknowledgements.....	8

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](#) [RFC 2119].

[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 OSPF [[RFC2328](#)] or IS-IS [ISO.10589.1992], i.e. all links in the domain MAY 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 L2 or L3 VPN scenario for example, a given 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 along the IP shortest path from one PE router to the other, all the links 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 it 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 distributed all the label bindings.

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

[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 in 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 if possible. In the case of OSPF, this cost is LSInfinity (16-bit value 0xFFFF) as proposed in [\[RFC3137\]](#). In the case of ISIS, the max metric value is $2^{24}-2$ (0xFFFFFE). Indeed, if a link is configured with $2^{24}-1$ (the maximum link metric per [\[RFC3784\]](#)) then this link is not advertised in the topology. It is important to keep the link in the topology to allow for 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

M. Jork, Alia Atlas, and L. Fang

[Page 3]

LDP IGP Synchronization

November 2008

condition can not generally be verified by a router and some estimated 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 the concerns of the large variations of the LIB table size and the equipment capabilities. In addition, this is a current work in progress on LDP End-of-LIB as specified in [LDP End-of-LIB], it enables the LDP speaker to signal the completion of its initial advertisement following session establish. 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](#)].

[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 is 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 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 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 which 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 any 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 link into the IGP, this becomes a local issue of the tunnel headend router.

[5.](#) Security Considerations

A DoS attack brings down LDP service on a link or prevents it from becoming operational on a link could be one of the possibilities that causes LDP related traffic blackholing. This document does not

address how to prevent LDP session failure. The mechanism described here is to prevent the link to be used when LDP is not operational while IGP is. Assigning the IGP cost to maximum on the link where LDP is failed and IGP is not should not introduce new security threats. The operation is internal in the router to allow LDP and IGP to communicate and react. Making many LDP links unavailable, however, is a security threat which can cause traffic being dropped due to limited available network capacity. This may be triggered by operational error or implementation error. They are considered as general Security issues and should follow the current best security practice.

[6.](#) IANA Considerations

M. Jork, Alia Atlas, and L. Fang

[Page 5]

LDP IGP Synchronization

November 2008

This document has no actions for IANA.

[7.](#) Normative References

[RFC5036] Andersson, L., Doolan, P., Feldman, N., Fredette, A., and B. Thomas, "LDP Specification", [RFC 5036](#), October 2007.

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

[8.](#) Informational References

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

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

[RFC3784] Smit, H., Li, T., Intermediate System to Intermediate System (IS-IS) Extension for Traffic Engineering, June 2004.

[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] Fang, L., Atlas, A., Chiussi, F., Kompella, K., and

Swallow, G., "LDP Failure Detection and Recovery", IEEE Communications Magazine, Vol.42, No.10, October 2004.

[LDP End-of-LIB] Asati, R., LDP End-of-LIB, [draft-ietf-mpls-ldp-end-of-lib-01.txt](#), September 2008.

9. Author's Addresses

Markus Jork
NextPoint Networks
3 Fedral St.
Billerica, MA 01821
USA
Email: mjork@nextpointnetworks.com

Alia Atlas
British Telecom
Email: alia.atlas@bt.com

M. Jork, Alia Atlas, and L. Fang

[Page 6]

LDP IGP Synchronization

November 2008

Luyuan Fang
Cisco Systems, Inc.
300 Beaver Brook Road
Boxborough, MA 01719
USA
Email: lufang@cisco.com

Full Copyright Statement

Copyright (C) The IETF Trust (2008).

This document is subject to the rights, licenses and restrictions contained in [BCP 78](#), and except as set forth therein, the authors retain all their rights.

This document and the information contained herein are provided on an "AS IS" basis and THE CONTRIBUTOR, THE ORGANIZATION HE/SHE REPRESENTS OR IS SPONSORED BY (IF ANY), THE INTERNET SOCIETY, THE IETF TRUST AND THE INTERNET ENGINEERING TASK FORCE DISCLAIM ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Intellectual Property

The IETF takes no position regarding the validity or scope of any Intellectual Property Rights or other rights that might be claimed to pertain to the implementation or use of the technology described in this document or the extent to which any license under such rights might or might not be available; nor does it represent that it has made any independent effort to identify any such rights. Information on the procedures with respect to rights in RFC documents can be found in [BCP 78](#) and [BCP 79](#).

Copies of IPR disclosures made to the IETF Secretariat and any assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of such proprietary rights by implementers or users of this specification can be obtained from the IETF on-line IPR repository at <http://www.ietf.org/ipr>.

The IETF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights that may cover technology that may be required to implement this standard. Please address the information to the IETF at ietf-ipr@ietf.org.

M. Jork, Alia Atlas, and L. Fang

[Page 7]

[10.](#) Acknowledgements

Funding for the RFC Editor function is provided by the IETF Administrative Support Activity (IASA).

The authors would like to thank Loa Andersson for his review and comments, thank Bruno Decraene for his in depth discussion, comments and helpful suggestions, and thank Dave Ward for his AD review comments and suggestions.

