Network Working Group B Decraene
Internet Draft JL Le Roux
Document: draft-decraene-mpls-ldp-interarea-00.txt France Telecom
Expiration Date: January 2006 July 2005

LDP extensions for Inter-Area LSP

Status of this Memo

Abstract

To facilitate the establishment of Label Switched Paths (LSP) that would span multiple IGP areas in a given autonomous system, this document proposes a new optional label mapping procedure for the Label Distribution Protocol (LDP).

This procedure allows the use of a label if the Forwarding Equivalence Class (FEC) Element matches an entry in the routing table (RIB). Matching is defined by an IP longest match search and does not mandate an exact match.

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 RFC-2119 [1].

2. Introduction

Link state IGP such as OSPF [RFC 2328] and IS-IS [RFC-1195] allow the partition of an autonomous system into areas or levels to increase routing scalability within a domain.

However, LDP mandates ([LDP] section 3.5.7.1 "Label Mapping Messages Procedures") that the IP address of the FEC Element *exactly* match an entry in the IP RIB. Therefore, the establishment of MPLS LSPs between LERs across areas/levels requires the redistribution of the exact (/32 for IPv4) loopback addresses of all the LERs across all areas.

As a consequence, the potential benefits that a multi-area domain may yield are diminished since the number of IP entries in the LSDB, RIB and FIB maintained by every LSR of the domain (whatever the area/level it belongs to) cannot be minimized.

Note however that IP prefixes and IGP events may still be reduced since IP addresses of links are usually not redistributed outside of their area.

The new label mapping procedure described in this document allows IP prefixes aggregation on the Area Border Routers and therefore eases the use of LDP in a multi-area/level context and restores the advantages introduced by IGP hierarchy.

3. Label Mapping Procedure

3.1. Control plane behavior

This document defines a new label mapping procedure for LDP. It MUST be possible to activate/deactivate this procedure by configuration.

With this new label mapping procedure, a LSR receiving a Label Mapping message from a neighbor LSR for a Prefix Address FEC Element SHOULD use the label for MPLS forwarding if its routing table contains an entry that matches the FEC Element and the advertising LSR is the next hop to reach the FEC. If so, it SHOULD advertise the FEC Element and a label to its LDP peers.

By "matching FEC Element", one should understand an IP longest match.

Note that with this new Label Mapping Procedure, each LSP established by LDP still strictly follows the shortest path(s) defined by the IGP.

Note also that this label mapping procedure only applies to an ordered label distribution control mode.

3.2. Populating Forwarding Tables

The new label mapping procedure MUST only populates the Next Hop Label Forwarding Entry (NHLFE) and the Incoming Label Map (ILM) defined in [MPLS] and sometimes referred to as Label Forwarding Information Base (LFIB).

This allows the establishment of multi-area LDP LSPs, without requiring any leaking of /32 loopback addresses across area boundaries.

The new label mapping procedure MUST neither populate the IP Forwarding Information Base (FIB) nor the FEC-to-NHLFE Map (FTN) defined in [MPLS]. FTN MUST only be populated by procedures described in [LDP] where the FEC exactly match an entry in the routing table.

Any protocol using LSP SHOULD be aware of these inter-area LSPs and SHOULD treat them like any other LSP. In particular (MP-)BGP SHOULD consider these LSP to route prefixes advertised by a BGP next-hop whose address correspond to the LSP FEC.

4. Application examples

4.1. Inter-area LSPs

Consider the following example of an autonomous system with one backbone area and two edge areas:

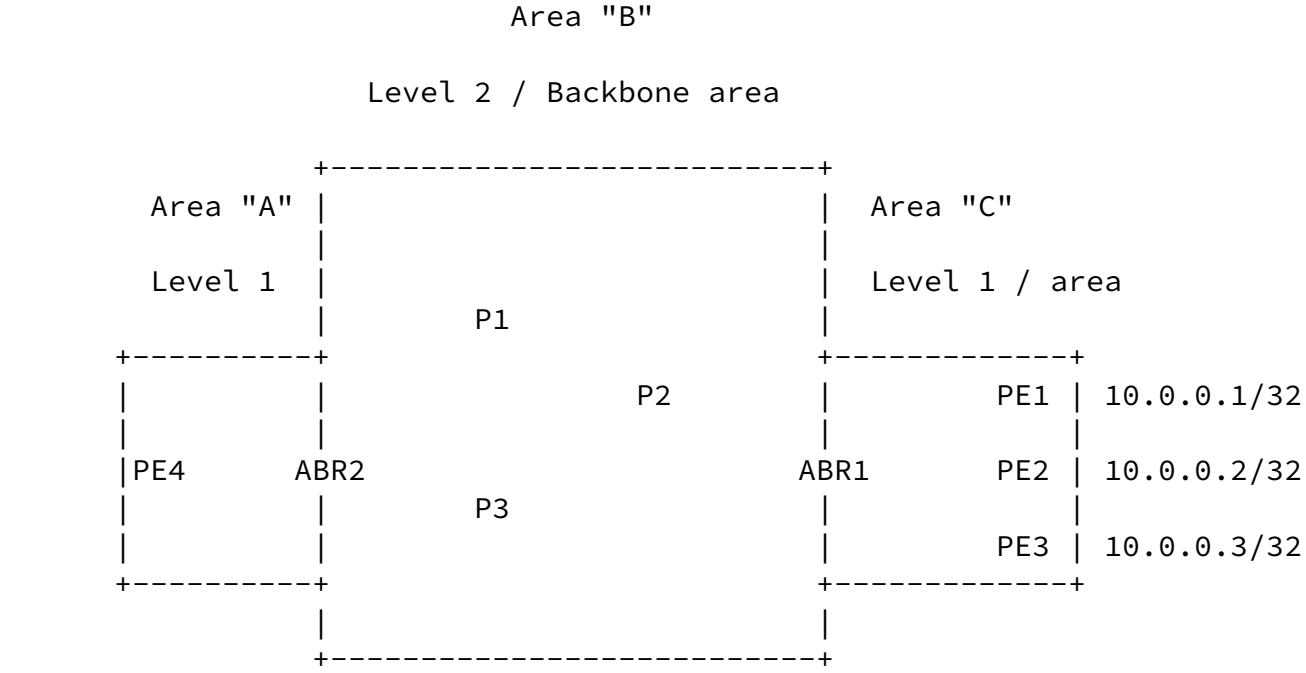

Figure 1: An IGP domain with two areas attached to the Backbone

Area.

Note that this applies equally to ISIS and OSPF. An ABR refers here either to an OSPF ABR or to an ISIS L1/L2 node.

All routers are MPLS enabled and MPLS connectivity is required between all PE routers.

In the "egress" area "C", the records available are:

IGP RIB	LDP FEC elements:
10.0.0.1/32	10.0.0.1/32
10.0.0.2/32	10.0.0.2/32
10.0.0.3/32	10.0.0.3/32

The area border router ABR1 advertises in the backbone area:
- the aggregated IP prefix 10.0.0.0/24 in the IGP
- all the individual IP FEC elements (/32) in LDP

In "backbone" area "B", the records available are:

IGP RIB	LDP FEC elements:
10.0.0.0/24	10.0.0.1/32
	10.0.0.2/32
	10.0.0.3/32

The area border router ABR2 advertises in the area "A":
- a default IP route (implicit or explicit) in the IGP
- all the individual IP FEC elements (/32) in LDP

In the "ingress" area "A", the records available are:

IGP RIB	LDP FEC elements:
0.0.0.0/0	10.0.0.1/32
	10.0.0.2/32
	10.0.0.3/32

In this situation, one LSP is established between ingress PE4 and every egress PE of area C.

4.2. Use of static routes

Consider the following example where a LER is dual connected to two LSRs:

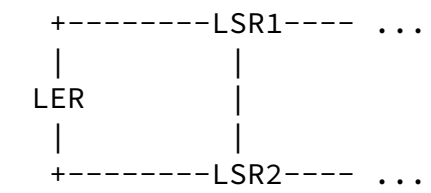

Figure 2: LER dual connected to two LSRs.

In some situations, especially on the edge of the network, it is valid to use static IP routes between the LER and the two LSRs. If necessary, the BFD protocol can be used to check the connectivity.

The current [LDP] specification would require the configuration and the maintenance of one IP route per egress LER and per outgoing interface.

The behavior described in this document would only require one IP route per outgoing interface.

5. Caveats for deployment

5.1. Deployment consideration

LSRs compliant with this document are backward compatible with LSRs that comply with [LDP].

For the successful establishment of end to end MPLS LSPs whose FEC are aggregated in the RIB, this new behavior must be implemented on all LSR in all areas where IP aggregation is used.

If all IP prefixes are leaked in the backbone area and only stub areas use IP aggregation, LSRs in the backbone area don't need to be compliant with this document.

5.2. Impact on routing convergence time

Introducing IGP Areas have a mixed impact regarding convergence time.

On one hand, it may increase convergence time as the ABR needs to process IGP topology messages and generates new ones. This is likely to require more time on the ABRs than simply flooding IGP messages. On the other hand, it reduces the network graph size in the IGP database of each router thus decrease computing time.

Performing IP aggregation on the ABR:
- requires careful IP addresses allocations to allow aggregation;
- can lead to sub-optimal routing;
- improves converge time as less entries need to be updated in the FIB.

In case of an egress LER failure in an area, the new procedure described in this document will increase the convergence time. As IP aggregation is performed, the IGP will not advertise the failure in remote areas. Therefore, the ingress LER will have to wait until the reception of the LDP withdraw message. As LDP messages are processed on each node before being propagated, information propagation is likely to be slower than with IGP messages which are flooded.

All in all, the convergence time is expected to be improved for link and P (LSR) node failures because FIBs have fewer entries to update. In case of a LER failure, the impact on convergence time is likely to be implementation dependant (IGP versus LDP processing) and network dependant (diameter size versus number of IP prefixes removed).

6. Security Considerations

The new Label Mapping procedure described in this document does not introduce any change as far as the Security Consideration section of [LDP] is concerned.

7. Normative References

[LDP] L. Andersson, P. Doolan, N. Feldman, A. Fredette, B. Thomas, "LDP Specification", RFC 3036, January 2001

[MPLS] E. Rosen, A. Viswanathan, R. Callon, " Multiprotocol Label Switching Architecture", RFC 3031, January 2001

[MP-BGP] Bates, T., Rekhter, Y, Chandra, R. and D. Katz, "Multiprotocol Extensions for BGP-4", RFC 2858, June 2000.

[BGP L3 VPN] E. Rosen, Y. Rekhter, "BGP/MPLS VPNs", RFC 2547, March 1999

[3107] Y. Rekhter, E. Rosen, "Carrying Label Information in BGP-4", RFC 3107, May 2001

[OSPF] J. Moy, "OSPF Version 2", RFC 1583, March 1994

[IS-IS] R. Callon, "Use of OSI IS-IS for Routing in TCP/IP and Dual Environments", RFC 1195, December 1990

8. Informative Reference

[BGP L2 VPN] K. Kompella, Y. Rekhter "Virtual Private LAN Service", draft-ietf-l2vpn-vpls-bgp-05.txt April 8, 2005, work in progress.

9. Acknowledgments

Authors would like to thank Ina Minei, Stefano Previdi, Benoit Fondeviole, Gilles Bourdon and Christian Jacquenet for the useful discussions on this subject, their review and comments.

10. Author's Addresses

Bruno Decraene
France Telecom
38-40 rue du General Leclerc
92794 Issy Moulineaux cedex 9
France
bruno.decraene@francetelecom.com

Jean-Louis Le Roux
France Telecom
2, avenue Pierre-Marzin
22307 Lannion Cedex
France
jeanlouis.leroux@francetelecom.com

11. Intellectual Property Considerations

The IETF takes no position regarding the validity or scope of any intellectual property 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; neither does it represent that it has made any effort to identify any such rights. Information on the IETF's procedures with respect to rights in standards-track and standards-related documentation can be found in BCP-11. Copies of claims of rights made available for publication 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 implementors or users of this specification can be obtained from the IETF Secretariat.

The IETF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights which may cover technology that may be required to practice this standard. Please address the information to the IETF Executive Director.

Disclaimer of Validity

Full Copyright Statement