Network Working Group
Internet Draft
Document: draft-decraene-mpls-ldp-interarea-01.txt
Expiration Date: April 2006

Proposed Status: Informational

B Decraene
JL Le Roux
France Telecom

I Minei
Juniper Networks, Inc.

October 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 (AS), 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].

IGP Area: OSPF Area or IS-IS level

ABR: OSPF Area Border Router or ISIS L1/L2 router

## 2. Introduction

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

However, LDP requires that the IP address of the FEC Element should *exactly* match an entry in the IP RIB: according to [LDP] section 3.5.7.1 (Label Mapping Messages Procedures) "An LSR receiving a Label Mapping message from a downstream LSR for a Prefix or Host Address FEC Element should not use the label for forwarding unless its routing table contains an entry that exactly matches the FEC Element.

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.

In that context, this document extend the Label Mapping Procedure defined in LDP so as to support the setup of inter-area LSPs while maintaining IP prefixes aggregation on the ABRs. This basically consists of extending the Label Mapping procedure so as to allow for "Longest Match Based" Label Mapping.

## 3. Label Mapping Procedure

This document defines a new label mapping procedure for LDP. It MUST be possible to activate/deactivate this procedure by configuration and it SHOULD be deactivated by default. It MAY be possible to activate it on a per prefix basis.





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 a 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.

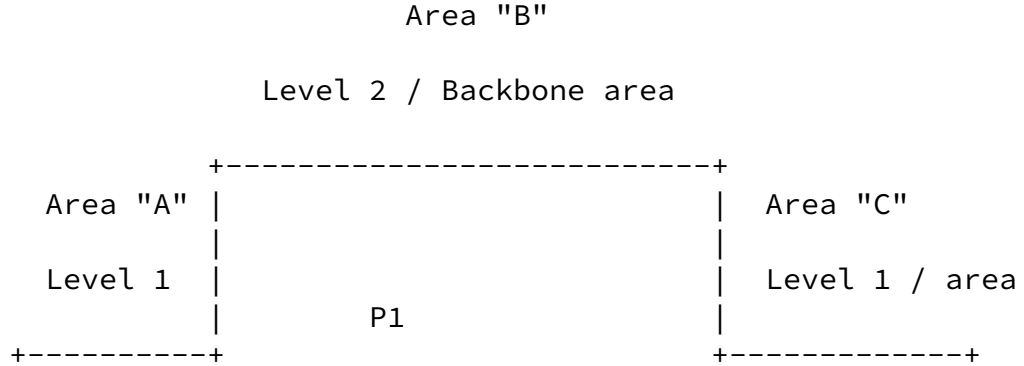
Note that with this new Label Mapping Procedure, each LSP established by LDP still strictly follows the shortest path(s) defined by the IGP.

FECs selected by this "Longest Match" label mapping procedure will be distributed in an ordered way. However this procedure is applicable to both independent and ordered distribution control mode.

## 4. Application examples

### 4.1. Inter-area LSPs

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



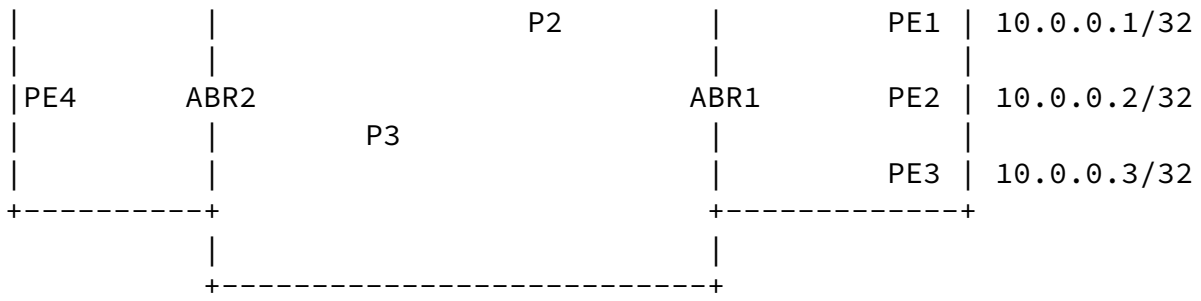



Figure 1: An IGP domain with two areas attached to the Backbone Area.

Note that this applies equally to IS-IS 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 (LSP) 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 10.0.0/24 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":
- an aggregated IP prefix 10.0/16 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:
  10.0/16                          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:





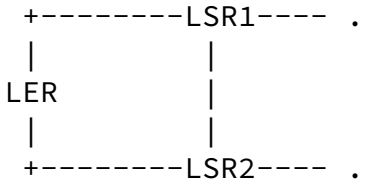



Figure 2: LER dual connected to two LSRs.

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 quickly detect loss of connectivity.

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

The new longest match Label Mapping Procedure 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

In case of an egress LER failure in an area, performing IP route aggregation on ABRs will change the routing convergence behavior. The IGP will not propagate the notification of the egress LER failure outside of the egress area and failure notification will rely on LDP signaling through the end to end propagation of the LDP withdraw message. This failure notification may be faster or longer depending on the implementations, the IGP timers used and the network topology (network diameter).

For link and P (LSR) node failures, the failure notification is unchanged and the convergence time is expected to be improved because RIB and FIBs have fewer entries to update.





## 6. Security Considerations

The "Longest Match" 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 Yakov Rekhter, 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

Ina Minei
Juniper Networks
1194 N. Mathilda Ave.
Sunnyvale, CA 94089
ina@juniper.net

## Intellectual Property Considerations



## Disclaimer of Validity

## Copyright Statement



## Acknowledgment

Funding for the RFC Editor function is currently provided by the Internet Society.