Network Working Group
Internet Draft

Expiration Date: November 2005

Naiming Shen
Enke Chen
Cisco Systems
Albert Tian
Redback Networks

Discovering LDP Next-Nexthop Labels

<draft-shen-mpls-ldp-nnhop-label-02.txt>

Status of this Memo

Abstract

This document specifies extensions to LDP in support of next-nexthop label discovery. The next-nexthop label information can be used to fast re-route LDP LSP traffic into an explicitly routed tunnel for nexthop node protection in the case of a link or node failure.

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 [5].

1. Introduction

As currently specified in [1], the LDP protocol only needs to know label mapping for the adjacent peers and there is no way for an LSR to learn the adjacent peer's downstream label mapping. This document proposes an LDP extension that allows an LSR to discover the next-nexthop label mapping from its downstream peers.

One application for learning the next-nexthop label mapping is for fast re-route. Similar to the facility based node-protection of LSP Fast ReRoute [2], the NFRR (Nexthop Fast ReRoute) [3] scheme allows an LSR to perform Fast ReRoute on any type of traffic, including LDP LSP traffic. When the Nexthop Fast ReRoute is used for node-protection of LDP LSP traffic, the next-nexthop labels are needed to tunnel the data traffic into the next-nexthop LSR in the case of a link or node failure.

A new Status TLV code is specified for an LSR to indicate its interest in receiving the next-nexthop label mapping information. A new Next-Nexthop Label TLV is specified to pass the downstream label mapping to the upstream LSR in the Label Mapping Message.

The extension specified in this document assumes the next-nexthop nodes use platform-wide label space for LDP. It is outside the scope of this document when the next-nexthop nodes use per-interface label space.

2. LDP Next-Nexthop Label Mapping Scheme

2.1 Example

Confiser LSRs interconnected with LDP as the following:

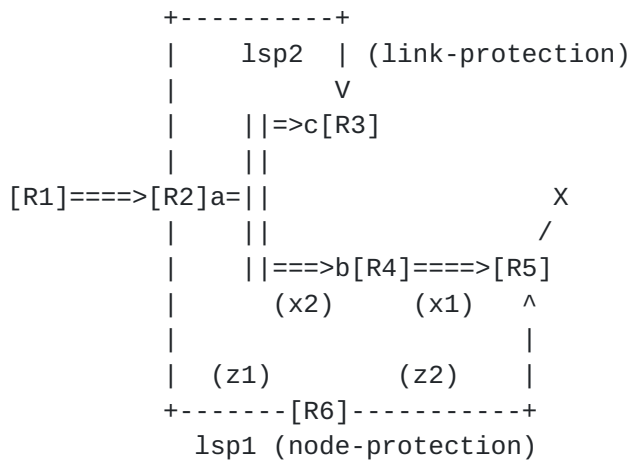


Figure 1: NFRR node-protection for LDP data traffic

R2 is the PLR (Point of Local Repair) node, the lsp1 is the

NFRR [3] LSP for the purpose of protecting node R4 over R2's interface "a". The lsp2 is the NFRR LSP for link protection in the case of interface "a" or "c" is down. We will only be concerned with node-protection using lsp1 in this document.

R5 advertises FEC X to R4 with label x1. R4 advertises the same FECs with label x2 to upstream peer R2. The RSVP signaled lsp1 uses label z1 from R2 to R6 and z2 from R6 to R5, and z2 can also be an implicit null.

When R2 detects either the interface "a" is down, or the nexthop "b" is unreachable, or LSR R4 is down, the forwarding engine on R2 will re-direct the LDP data traffic into the NFRR tunnel lsp1. This can be quickly done by pushing the label x1 onto the label stack and send the packet through the lsp1 for LDP data traffic going to FEC X. As long as the platform-wide label space is used on LSR R5, the R5 does not even know the difference. In this case, the next-nexthop label x1 is used by PLR node R2 for fast re-route with node-protection. For this scheme to work, LSR R4 needs to advertise the next-nexthop label x1 to the upstream LSR R2 in addition to their own label mapping of x2 for the same FEC.

2.2 Next-Nexthop Label Request

Take the same example as in section 2.1, a user can statically configure on LSR R4 that it needs to include downstream labels to all or some of the upstream peers while it advertises the label mappings. A better way is for LSR R2 to make a request to its peer R4 that it is interested in receiving the next-nexthop label mapping information, since R2 has already been configured to perform node-protection for LSR R4.

When the LDP peer between R2 and R4 is up, and there is at least one NFRR lsp configured on R2 to perform node-protection of R4, R2 can optionally send a Notification Message with the Next-Nexthop Label Request bit set in the Status TLV. When the last NFRR LSP protecting node R4 is removed, R2 can optionally send the Notification Message to R4 with the Next-Nexthop Label Withdraw bit set in the Status TLV.

2.3 Next-Nexthop Label Advertisement

When an LSR advertises the FEC-label bindings to its peer, if it has received the Next-Nexthop Label Request from that peer or the LSR is configured with this capability, it SHOULD include the next-nexthop label mapping information when applicable in the Label Mapping Message.

An optional Next-Nexthop Label TLV is defined to be used in the Label Mapping Message. The Next-Nexthop Label contains a list

of (label, downstream router-id) tuples. More than one tuple can be used when there is an ECMP case to different downstream nodes for the same FECs. It is an implementation and local configuration issue whether to announce only one or multiple tuples in the ECMP case.

If some FECs are not advertised with next-nexthop labels, then no node-protection can be performed on those FECs. But they can still be fast re-routed with NFRR link-protection scheme [3]. If there is a NFRR LSP built from R2 to R4, then the LDP data traffic will be re-routed directly onto R4 itself. The node-protection is not meant for all the situations. Usually node-protection is used in the backbone portion of the network, and link-protection is used close to the edge of the network.

2.4 Next-Nexthop Label Update

If an LSR advertises the Next-Nexthop Label TLV in the Label Mapping Messages, and when the next-nexthop label information changes, it MUST re-send the Label Mapping Message with updated next-nexthop label information. The LSR SHOULD implement a means to dampen the re-advertisement to avoid potentially excessive updating due to link flapping.

2.5 Comparing with Targeted LDP Session Approach

The discovering Next-Nexthop LDP label scheme described in this document relies on the downstream LDP nodes to relay the label mapping of Next-Nexthop LDP nodes. This approach has a number of advantages in comparing with directly setting up targeted LDP sessions to the Next-Nexthop LDP nodes.

When using the downstream LDP nodes relaying label mapping message, not only there is less configuration involved, but also the network will have less LDP sessions to be established.

In the case of two and more Next-Nexthop LDP nodes advertising the same route to their immediate upstream nodes, the selection of the nexthops follows the underlying routing. This is not the case in the targeted LDP session in general. When the PLR receives from two Next-Nexthop LDP nodes on the label binding of the same route directly, there is no way for it to tell which one should be preferred or both of them need to be used for loadsharing. The targeted LDP session approach might couple the IGP topology information in the route selection process, but it would make the solution more complex especially in the case of nexthop LDP node being an area border router.

3. Next-Nexthop Label Packet Encoding

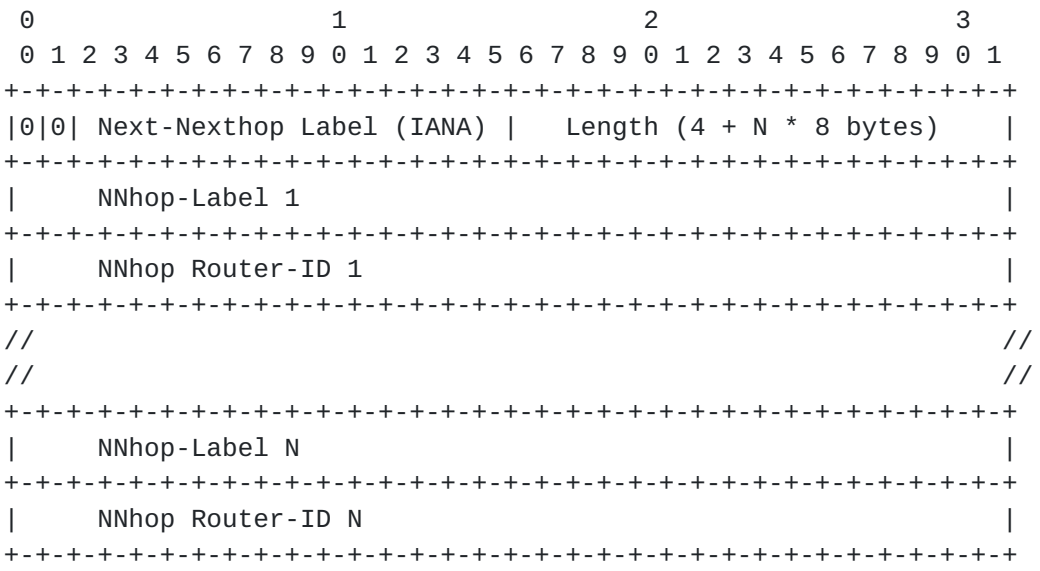
3.1 Next-Nexthop Label Bits in Status TLV

The Next-Nexthop Label Request/Withdraw information is sent in the Notification Message. Two bits (to be allocated by IANA) are defined in this document, one for Request and one for Withdraw. Unlike most of the bits already defined in the Status TLV, the Next-Nexthop Label Bits are used by an LSR to dynamically announce a capability to its peers.

The E bit and F bit MUST be set to zero if Next-Nexthop Label Request or Withdraw is the only status code set. The Next-Nexthop Label Bits SHOULD only be used in Notification Message, otherwise it MUST be quietly ignored upon receipt.

3.2 Next-Nexthop Label TLV in Label Mapping Message

The Next-Nexthop Label TLV can be optionally carried in the Optional Parameters field of a Label Mapping Message. The TLV consists a list of (label, router-id) pairs with the following format:



NNhop-Label
: Next-Nexthop Label. This is a 20-bit label value as specified in [4] represented as a 20-bit number in a 4 octet field.

NNhop Router-ID
: Next-Nexthop router-ID which advertised that next-nexthop label. This is a 4 octet number.

4. Security Considerations

This mechanism does not introduce any new security issue in LDP.

5. IANA Considerations

Two new bits in Status TLV and a new LDP TLV Type is defined in section 3. This LDP extension requires that IANA allocate those numbers.

6. Acknowledgments

TBD.

7. References

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

[2] Pan, P., Gan, D., Swallow, G., Vasseur, J.Ph., Copper, D., Atlas, A., Jork, M., "Fast Reroute Technique in RSVP-TE", Internet draft, draft-ietf-mpls-rsvp-lsp-fastreroute-07.txt, work in progress.

[3] Shen, N., Pan, P., "Nexthop Fast ReRoute for IP and MPLS", Internet draft, draft-shen-nhop-fastreroute-01.txt, work in progress.

[4] Rosen, E., Tappan, D., Federkow, G., Rekhter, Y., Farinacci, D., Li, T. and A. Conta, "MPLS Label Stack Encoding", RFC 3032, January 2001.

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

8. Authors' Addresses

Naiming Shen
Cisco Systems
170 W. Tasman Drive
San Jose, CA, 95134 USA
e-mail: naiming@cisco.com

Enke Chen
Cisco Systems
170 W. Tasman Drive
San Jose, CA, 95134 USA
e-mail: enke@cisco.com

Albert Tian
Redback Networks, Inc.
300 Holger Way
San Jose, CA 95134
e-mail: tian@redback.com

Intellectual Property

Full Copyright Notice

EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT
THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR

Acknowledgment

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