

SPRING Working Group
Internet-Draft
Intended status: Standards Track
Expires: October 28, 2020

G. Mirsky
ZTE Corp.
J. Tantsura
Apstra, Inc.
I. Varlashkin
Google
M. Chen
Huawei
J. Wenying
CMCC
April 26, 2020

Bidirectional Forwarding Detection (BFD) in Segment Routing Networks
Using MPLS Dataplane
draft-mirsky-spring-bfd-10

Abstract

Segment Routing (SR) architecture leverages the paradigm of source routing. It can be realized in the Multiprotocol Label Switching (MPLS) network without any change to the data plane. A segment is encoded as an MPLS label, and an ordered list of segments is encoded as a stack of labels. Bidirectional Forwarding Detection (BFD) is expected to monitor any existing path between systems. This document defines how to use Label Switched Path Ping to bootstrap a BFD session, control an SR Policy in the reverse direction of the SR-MPLS tunnel, and applicability of BFD Demand mode in the SR-MPLS domain. Also, the document describes the use of BFD Echo with BFD Control packet payload.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <https://datatracker.ietf.org/drafts/current/>.

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."

This Internet-Draft will expire on October 28, 2020.

Copyright Notice

Copyright (c) 2020 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<https://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction	2
1.1. Conventions	3
1.1.1. Terminology	3
1.1.2. Requirements Language	3
2. Bootstrapping BFD Session over Segment Routed Tunnel with MPLS Data Plane	4
3. Use BFD Reverse Path TLV over Segment Routed MPLS Tunnel	5
4. Use Non-FEC Path TLV	5
5. BFD Reverse Path TLV over Segment Routed MPLS Tunnel with Dynamic Control Plane	7
6. Applicability of BFD Demand Mode in SR-MPLS Domain	7
7. Using BFD to Monitor Point-to-Multipoint SR Policy	7
8. Use of Echo BFD in SR-MPLS	8
9. IANA Considerations	8
9.1. Non-FEC Path TLV	8
9.2. Return Code	9
10. Implementation Status	10
11. Security Considerations	10
12. Contributors	11
13. Acknowledgments	11
14. References	11
14.1. Normative References	11
14.2. Informative References	13
Authors' Addresses	13

1. Introduction

[RFC5880], [RFC5881], and [RFC5883] defined the operation of Bidirectional Forwarding Detection (BFD) protocol between the two systems over IP networks. [RFC5884] and [RFC7726] set rules for using BFD Asynchronous mode over point-to-point (p2p) Multiprotocol

Label Switching (MPLS) Label Switched Path (LSP). These latter standards implicitly assume that the remote BFD system, which is at the egress Label Edge Router (LER), will use the shortest path route regardless of the path the BFD system at the ingress LER uses to send BFD Control packets towards it. Throughout this document, references to ingress LER and egress LER are used, respectively, as a shortened version of the "BFD system at the ingress/egress LER".

This document defines the use of LSP Ping for Segment Routing networks over MPLS data plane [RFC8287] to bootstrap and control path of a BFD session from the egress to ingress LER using Segment Routing tunnel with MPLS data plane (SR-MPLS).

1.1. Conventions

1.1.1. Terminology

BFD: Bidirectional Forwarding Detection

BSID: Binding Segment Identifier

FEC: Forwarding Equivalence Class

MPLS: Multiprotocol Label Switching

SR-MPLS Segment Routing with MPLS data plane

LSP: Label Switched Path

LER Label Edge Router

p2p Point-to-point

p2mp Point-to-multipoint

SID Segment Identifier

SR Segment Routing

1.1.2. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

2. Bootstrapping BFD Session over Segment Routed Tunnel with MPLS Data Plane

Use of an LSP Ping to bootstrap BFD over MPLS LSP is required, as documented in [RFC5884], to establish an association between a fault detection message, i.e., BFD Control message, and the Forwarding Equivalency Class (FEC) of a single label stack LSP in case of Penultimate Hop Popping or when the egress LER distributes the Explicit NULL label to the penultimate hop router. The Explicit NULL label is not advertised as a Segment Identifier (SID) by an SR node but, as demonstrated in section 3.1 [RFC8660] if the operation at the penultimate hop is NEXT; then the egress SR node will receive an IP encapsulated packet. Thus the conclusion is that LSP Ping MUST be used to bootstrap a BFD session in an SR-MPLS domain if there are no other means to bootstrap the BFD session, e.g., using an extension to a dynamic routing protocol as described in [I-D.ietf-bess-mvpn-fast-failover] and [I-D.ietf-pim-bfd-p2mp-use-case].

As demonstrated in [RFC8287], the introduction of Segment Routing network domains with an MPLS data plane requires three new sub-TLVs that MAY be used with Target FEC TLV. Section 6.1 addresses the use of the new sub-TLVs in Target FEC TLV in LSP ping and LSP traceroute. For the case of LSP ping, the [RFC8287] states that:

The initiator, i.e., ingress LER, MUST include FEC(s) corresponding to the destination segment.

The initiator MAY include FECs corresponding to some or all of segments imposed in the label stack by the ingress LER to communicate the segments traversed.

It has been noted in [RFC5884] that a BFD session monitors for defects particular <MPLS LSP, FEC> tuple. [RFC7726] clarified how to establish and operate multiple BFD sessions for the same <MPLS LSP, FEC> tuple. Because only the ingress LER is aware of the SR-based explicit route, the egress LER can associate the LSP ping with BFD Discriminator TLV with only one of the FECs it advertised for the particular segment. Thus this document clarifies that:

When LSP Ping is used to bootstrapping a BFD session for SR-MPLS tunnel the FEC corresponding to the segment to be associated with the BFD session MUST be as the very last sub-TLV in the Target FEC TLV.

If the target segment is an anycast prefix segment ([I-D.ietf-spring-mpls-anycast-segments]) the corresponding Anycast SID MUST be included in the Target TLV as the very last sub-TLV.

Also, for BFD Control packet the ingress SR node MUST use precisely the same label stack encapsulation, especially Entropy Label ([RFC6790]), as for the LSP ping with the BFD Discriminator TLV that bootstrapped the BFD session. Other operational aspects of using BFD to monitor the continuity of the path to the particular Anycast SID, advertised by a group of SR-MPLS capable nodes, will be considered in the future versions of the document.

Encapsulation of a BFD Control packet in Segment Routing network with MPLS data plane MUST follow Section 7 [RFC5884] when the IP/UDP header used and MUST follow Section 3.4 [RFC6428] without IP/UDP header being used.

3. Use BFD Reverse Path TLV over Segment Routed MPLS Tunnel

For BFD over MPLS LSP case, per [RFC5884], egress LER MAY send BFD Control packet to the ingress LER either over IP network or an MPLS LSP. Similarly, for the case of BFD over p2p SR-MPLS tunnel, the egress LER MAY route BFD Control packet over the IP network, as described in [RFC5883], or transmit over a segment tunnel, as described in Section 7 [RFC5884]. In some cases, there may be a need to direct egress LER to use a specific path for the reverse direction of the BFD session by using the BFD Reverse Path TLV and following all procedures as defined in [I-D.ietf-mpls-bfd-directed].

4. Use Non-FEC Path TLV

For the case of MPLS data plane, Segment Routing Architecture [RFC8402] explains that "a segment is encoded as an MPLS label. An ordered list of segments is encoded as a stack of labels."

This document defines a new optional Non-FEC Path TLV. The format of the Non-FEC Path TLV is presented in Figure 1

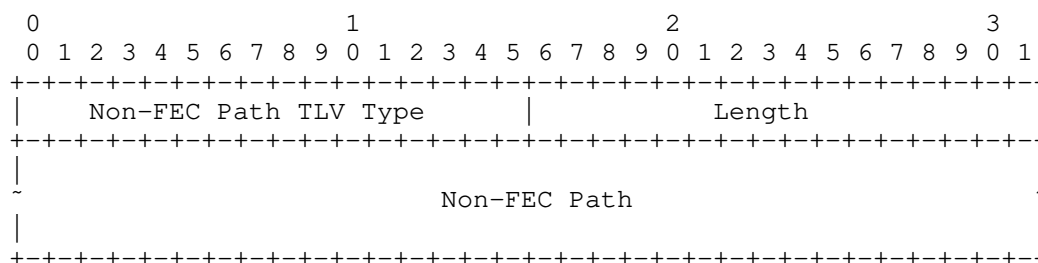


Figure 1: Non-FEC Path TLV Format

Non-FEC Path TLV Type is two octets in length and has a value of TBD1 (to be assigned by IANA as requested in Section 9.1).

Length field is two octets long and defines the length in octets of the Non-FEC Path field.

Non-FEC Path field contains a sub-TLV. Any Non-FEC Path sub-TLV (defined in this document or to be defined in the future) for Non-FEC Path TLV type MAY be used in this field. None or one sub-TLV MAY be included in the Non-FEC Path TLV. If no sub-TLV has been found in the Non-FEC Path TLV, the egress LER MUST revert to using the reverse path selected based on its local policy. If there is more than one sub-TLV, then the Return Code in echo reply MUST be set to value TBD3 "Too Many TLVs Detected" (to be assigned by IANA as requested in Table 4).

Non-FEC Path TLV MAY be used to specify the reverse path of the BFD session identified in the BFD Discriminator TLV. If the Non-FEC Path TLV is present in the echo request message the BFD Discriminator TLV MUST be present as well. If the BFD Discriminator TLV is absent when the Non-FEC Path TLV is included, then it MUST be treated as malformed Echo Request, as described in [RFC8029].

This document defines the Segment Routing MPLS Tunnel sub-TLV that MAY be used with the Non-FEC Path TLV. The format of the sub-TLV is presented in Figure 2.

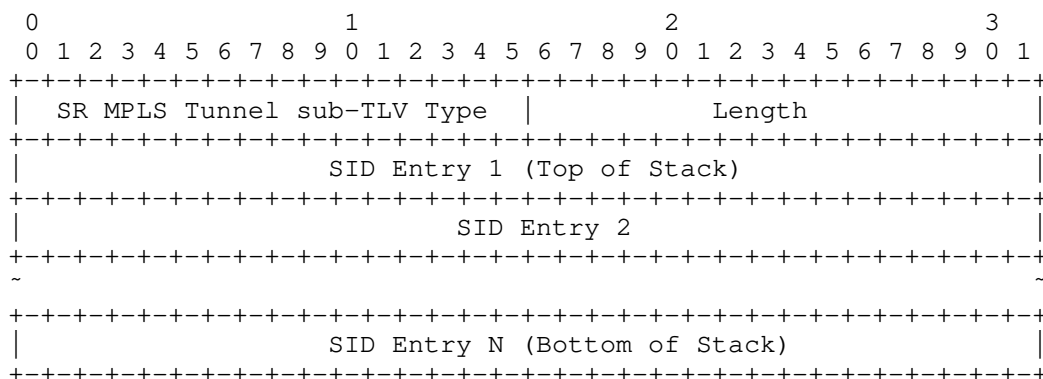


Figure 2: Segment Routing MPLS Tunnel sub-TLV

The Segment Routing MPLS Tunnel sub-TLV Type is two octets in length, and has a value of TBD2 (to be assigned by IANA as requested in Section 9.1).

The egress LER MUST use the Value field as label stack for BFD Control packets for the BFD session identified by the source IP

address of the MPLS LSP Ping packet and the value in the BFD Discriminator TLV. Label Entries MUST be in network order.

5. BFD Reverse Path TLV over Segment Routed MPLS Tunnel with Dynamic Control Plane

When Segment Routed domain with MPLS data plane uses distributed tunnel computation BFD Reverse Path TLV MAY use Target FEC sub-TLVs defined in [RFC8287].

6. Applicability of BFD Demand Mode in SR-MPLS Domain

[I-D.mirsky-bfd-mpls-demand] defines how Demand mode of BFD, specified in sections 6.6 and 6.18.4 of [RFC5880], can be used to monitor uni-directional MPLS LSP. Similar procedures can be following in SR-MPLS to monitor uni-directional SR tunnels:

- o an ingress SR node bootstraps BFD session over SR-MPLS in Async BFD mode;
- o once BFD session is Up, the ingress SR node switches the egress LER into the Demand mode by setting D field in BFD Control packet it transmits;
- o if the egress LER detects the failure of the BFD session, it sends its BFD Control packet to the ingress SR node over the IP network with a Poll sequence;
- o if the ingress SR node receives a BFD Control packet from the remote node in a Demand mode with Poll sequence and Diag field indicating the failure, the ingress SR node transmits BFD Control packet with Final over IP and switches the BFD over SR-MPLS back into Async mode, sending BFD Control packets one per second.

7. Using BFD to Monitor Point-to-Multipoint SR Policy

[I-D.voyer-spring-sr-p2mp-policy] defined variants of SR Policy to deliver point-to-multipoint (p2mp) services. For the given P2MP segment [RFC8562] can be used if, for example, leaves have an alternative source of the multicast service flow to select. In such a scenario, a leaf may switch to using the alternative flow after p2mp BFD detects the failure in the working multicast path. For scenarios where it is required for the root to monitor the state of the multicast tree [RFC8563] can be used. The root may use the detection of the failure of the multicast tree to the particular leaf to restore the path for that leaf or re-instantiate the whole multicast tree.

An essential part of using p2mp BFD is the bootstrapping the BFD session at all the leaves. The root, acting as the MultipointHead, MAY use LSP Ping with the BFD Discriminator TLV. Alternatively, extensions to routing protocols, e.g., BGP, or management plane, e.g., PCEP, MAY be used to associate the particular P2MP segment with MultipointHead's Discriminator. Extensions for routing protocols and management plane are for further study.

8. Use of Echo BFD in SR-MPLS

Echo-BFD [RFC5880] can be used to monitor an SR Policy between the local and the remote BFD peers. As defined in [RFC5880], the remote BFD system does not process the payload of an Echo BFD. Thus it is the local system that demultiplexes the Echo BFD packet matching it to the appropriate BFD session and detects missing Echo BFD packets. A BFD Control packet MAY be used as the payload of Echo BFD. This specification defines the use of Echo BFD in SR-MPLS network with BFD Control packet as the payload. The use of other types of Echo BFD payload is outside the scope of this document. Because the remote BFD system does not process Echo BFD, the value of the Your Discriminator field MUST be set to the discriminator the local BFD system assigned to the given BFD session. My Discriminator field MUST be zeroed. Authentication MUST be set according to the configuration of the BFD session. To ensure that the Echo BFD packet is returned to the sender without being processed, the sender MAY use a Binding SID (BSID) [RFC8402] that has been bound with the SR Policy that ensures the return of a packet to that particular node. A BSID MAY be associated with the SR Policy that is the reverse to the SR Policy programmed onto the BFD Echo packet by the sender.

9. IANA Considerations

9.1. Non-FEC Path TLV

IANA is requested to assign new TLV type from the from Standards Action range of the registry "Multiprotocol Label Switching Architecture (MPLS) Label Switched Paths (LSPs) Ping Parameters - TLVs" as defined in Table 1.

Value	TLV Name	Reference
TBD1	Non-FEC Path TLV	This document

Table 1: New Non-FEC Path TLV

IANA is requested to create new Non-FEC Path sub-TLV registry for the Non-FEC Path TLV, as described in Table 2.

Range	Registration Procedures	Note
0-16383	Standards Action	This range is for mandatory TLVs or for optional TLVs that require an error message if not recognized. Experimental RFC needed
16384-31743	Specification Required	
32768-49161	Standards Action	This range is for optional TLVs that can be silently dropped if not recognized. Experimental RFC needed
49162-64511	Specification Required	
64512-65535	Private Use	

Table 2: Non-FEC Path sub-TLV registry

IANA is requested to allocate the following values from the Non-FEC Path sub-TLV registry as defined in Table 3.

Value	Description	Reference
0	Reserved	This document
TBD2	Segment Routing MPLS Tunnel sub-TLV	This document
65535	Reserved	This document

Table 3: New Segment Routing Tunnel sub-TLV

9.2. Return Code

IANA is requested to create Non-FEC Path sub-TLV sub-registry for the new Non-FEC Path TLV and assign a new Return Code value from the "Multi-Protocol Label Switching (MPLS) Label Switched Paths (LSPs) Ping Parameters" registry, "Return Codes" sub-registry, as follows using a Standards Action value.

Value	Description	Reference
X TBD3	Too Many TLVs Detected.	This document

Table 4: New Return Code

10. Implementation Status

- The organization responsible for the implementation: ZTE Corporation.
- The implementation's name ROSng SW empowers traditional routers, e.g., ZXCTN 6000.
- A brief general description: A list of SIDs can be specified as the Return Path for an SR-MPLS tunnel.
- The implementation's level of maturity: production.
- Coverage: complete
- Version compatibility: draft-mirsky-spring-bfd-06.
- Licensing: proprietary.
- Implementation experience: Appreciate Early Allocation of values for Non-FEC TLV and Segment Routing MPLS Tunnel sub-TLV (using Private Use code points).
- Contact information: Qian Xin qian.xin2@zte.com.cn
- The date when information about this particular implementation was last updated: 12/16/2019

Note to RFC Editor: This section MUST be removed before publication of the document.

11. Security Considerations

Security considerations discussed in [RFC5880], [RFC5884], [RFC7726], and [RFC8029] apply to this document.

12. Contributors

Xiao Min
ZTE Corp.
Email: xiao.min2@zte.com.cn

13. Acknowledgments

Authors greatly appreciate the help of Qian Xin, who provided the information about the implementation of this specification.

14. References

14.1. Normative References

- [I-D.ietf-mpls-bfd-directed]
Mirsky, G., Tantsura, J., Varlashkin, I., and M. Chen,
"Bidirectional Forwarding Detection (BFD) Directed Return
Path", draft-ietf-mpls-bfd-directed-13 (work in progress),
December 2019.
- [I-D.mirsky-bfd-mpls-demand]
Mirsky, G., "BFD in Demand Mode over Point-to-Point MPLS
LSP", draft-mirsky-bfd-mpls-demand-06 (work in progress),
December 2019.
- [I-D.voyer-spring-sr-p2mp-policy]
daniel.voyer@bell.ca, d., Filsfils, C., Parekh, R.,
Bidgoli, H., and Z. Zhang, "SR Replication Policy for P2MP
Service Delivery", draft-voyer-spring-sr-p2mp-policy-03
(work in progress), July 2019.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate
Requirement Levels", BCP 14, RFC 2119,
DOI 10.17487/RFC2119, March 1997,
<<https://www.rfc-editor.org/info/rfc2119>>.
- [RFC5880] Katz, D. and D. Ward, "Bidirectional Forwarding Detection
(BFD)", RFC 5880, DOI 10.17487/RFC5880, June 2010,
<<https://www.rfc-editor.org/info/rfc5880>>.
- [RFC5881] Katz, D. and D. Ward, "Bidirectional Forwarding Detection
(BFD) for IPv4 and IPv6 (Single Hop)", RFC 5881,
DOI 10.17487/RFC5881, June 2010,
<<https://www.rfc-editor.org/info/rfc5881>>.

- [RFC5883] Katz, D. and D. Ward, "Bidirectional Forwarding Detection (BFD) for Multihop Paths", RFC 5883, DOI 10.17487/RFC5883, June 2010, <<https://www.rfc-editor.org/info/rfc5883>>.
- [RFC5884] Aggarwal, R., Kompella, K., Nadeau, T., and G. Swallow, "Bidirectional Forwarding Detection (BFD) for MPLS Label Switched Paths (LSPs)", RFC 5884, DOI 10.17487/RFC5884, June 2010, <<https://www.rfc-editor.org/info/rfc5884>>.
- [RFC6428] Allan, D., Ed., Swallow, G., Ed., and J. Drake, Ed., "Proactive Connectivity Verification, Continuity Check, and Remote Defect Indication for the MPLS Transport Profile", RFC 6428, DOI 10.17487/RFC6428, November 2011, <<https://www.rfc-editor.org/info/rfc6428>>.
- [RFC7726] Govindan, V., Rajaraman, K., Mirsky, G., Akiya, N., and S. Aldrin, "Clarifying Procedures for Establishing BFD Sessions for MPLS Label Switched Paths (LSPs)", RFC 7726, DOI 10.17487/RFC7726, January 2016, <<https://www.rfc-editor.org/info/rfc7726>>.
- [RFC8029] Kompella, K., Swallow, G., Pignataro, C., Ed., Kumar, N., Aldrin, S., and M. Chen, "Detecting Multiprotocol Label Switched (MPLS) Data-Plane Failures", RFC 8029, DOI 10.17487/RFC8029, March 2017, <<https://www.rfc-editor.org/info/rfc8029>>.
- [RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <<https://www.rfc-editor.org/info/rfc8174>>.
- [RFC8287] Kumar, N., Ed., Pignataro, C., Ed., Swallow, G., Akiya, N., Kini, S., and M. Chen, "Label Switched Path (LSP) Ping/Traceroute for Segment Routing (SR) IGP-Prefix and IGP-Adjacency Segment Identifiers (SIDs) with MPLS Data Planes", RFC 8287, DOI 10.17487/RFC8287, December 2017, <<https://www.rfc-editor.org/info/rfc8287>>.
- [RFC8402] Filss, C., Ed., Previdi, S., Ed., Ginsberg, L., Decraene, B., Litkowski, S., and R. Shakir, "Segment Routing Architecture", RFC 8402, DOI 10.17487/RFC8402, July 2018, <<https://www.rfc-editor.org/info/rfc8402>>.
- [RFC8562] Katz, D., Ward, D., Pallagatti, S., Ed., and G. Mirsky, Ed., "Bidirectional Forwarding Detection (BFD) for Multipoint Networks", RFC 8562, DOI 10.17487/RFC8562, April 2019, <<https://www.rfc-editor.org/info/rfc8562>>.

- [RFC8563] Katz, D., Ward, D., Pallagatti, S., Ed., and G. Mirsky, Ed., "Bidirectional Forwarding Detection (BFD) Multipoint Active Tails", RFC 8563, DOI 10.17487/RFC8563, April 2019, <<https://www.rfc-editor.org/info/rfc8563>>.
- [RFC8660] Bashandy, A., Ed., Filsfils, C., Ed., Previdi, S., Decraene, B., Litkowski, S., and R. Shakir, "Segment Routing with the MPLS Data Plane", RFC 8660, DOI 10.17487/RFC8660, December 2019, <<https://www.rfc-editor.org/info/rfc8660>>.

14.2. Informative References

- [I-D.ietf-bess-mvpn-fast-failover]
Morin, T., Kebler, R., and G. Mirsky, "Multicast VPN fast upstream failover", draft-ietf-bess-mvpn-fast-failover-10 (work in progress), February 2020.
- [I-D.ietf-pim-bfd-p2mp-use-case]
Mirsky, G. and J. Xiaoli, "Bidirectional Forwarding Detection (BFD) for Multi-point Networks and Protocol Independent Multicast - Sparse Mode (PIM-SM) Use Case", draft-ietf-pim-bfd-p2mp-use-case-03 (work in progress), January 2020.
- [I-D.ietf-spring-mpls-anycast-segments]
Sarkar, P., Gredler, H., Filsfils, C., Previdi, S., Decraene, B., and M. Horneffer, "Anycast Segments in MPLS based Segment Routing", draft-ietf-spring-mpls-anycast-segments-02 (work in progress), January 2018.
- [RFC6790] Kompella, K., Drake, J., Amante, S., Henderickx, W., and L. Yong, "The Use of Entropy Labels in MPLS Forwarding", RFC 6790, DOI 10.17487/RFC6790, November 2012, <<https://www.rfc-editor.org/info/rfc6790>>.

Authors' Addresses

Greg Mirsky
ZTE Corp.

Email: gregimirsky@gmail.com

Jeff Tantsura
Apstra, Inc.

Email: jefftant.ietf@gmail.com

Ilya Varlashkin
Google

Email: Ilya@nobulus.com

Mach (Guoyi) Chen
Huawei

Email: mach.chen@huawei.com

Jiang Wenying
CMCC

Email: jiangwenying@chinamobile.com