Network Working Group
Internet Draft
Intended status: Standards Track
Expires: August 31, 2024

Y. Liu
W. Cheng
China Mobile
C. Lin
M. Chen
New H3C Technologies
X. Min
ZTE

March 4, 2024

Encapsulation of BFD for SRv6 Policy
draft-liu-spring-bfd-srv6-policy-encap-03

Abstract

Bidirectional Forwarding Detection (BFD) mechanisms can be used for fast detection of failures in the forwarding path of SR Policy. This document describes the encapsulation of BFD for SRv6 Policy. The BFD packets may be encapsulated in Insert-mode or Encaps-mode.

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 August 31, 2024.

Table of Contents

1. Introduction

Segment Routing (SR) [RFC8402] allows a headend node to steer a packet flow along any path. Per-path states of Intermediate nodes are eliminated thanks to source routing. A Segment Routing Policy (SR Policy) [RFC9256] is an ordered list of segments (i.e., instructions) that represent a source-routed policy. The packets steered into an SR Policy carry an ordered list of segments associated with that SR Policy. The SRv6 Policy is the instantiation of SR Policy for SR over IPv6 (SRv6) data plane.

In order to provide end-to-end protection, the headend node need to rapidly detect any failures in the forwarding path of SR Policy, so that it could switch from the active candidate path to another backup candidate path within the same SR Policy or switch from the active SR Policy to another backup SR Policy. Bidirectional Forwarding Detection (BFD) mechanisms [RFC5880] [RFC7880] can be used for fast failure detection of P2P SR Policy.

[RFC8562] defines a method of using BFD to monitor and detect unicast failures between a sender and multipoint receivers, which

can be used for fast failure detection of P2MP SRv6 Policy [I-D.ietf-pim-sr-p2mp-policy].

As specified in [I-D.draft-ietf-spring-bfd], the basic element monitored by the BFD is a segment list that is a constituent of the candidate path of the particular SR Policy.

An SR Policy may consist of multiple candidate paths, and each candidate path may consist of multiple segment lists. When the associated BFD session is failed, a segment list becomes invalid. If some of the segment lists fail, the forwarding will be weighted load-balancing among the other segment lists. If all of the segment lists fail, the candidate path becomes invalid. If the active candidate path fails, the switchover to another backup candidate path will be triggered. If all the candidate paths fail, the SR Policy becomes invalid.

This document describes the encapsulation of BFD [RFC5880] [RFC7880] [RFC8562] for SRv6 Policy. BFD Demand Mode and BFD Echo Function are out of the scope of this document.

1.1. 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. Encapsulation of BFD Packet for SRv6 Policy

On SRv6 data plane, a BFD packet for a segment list of an SRv6 Policy carries a Segment Routing Header (SRH) [RFC8754] containing a list of SRv6 SIDs associated with that segment list.

BFD packets may be encapsulated in Insert-mode or Encaps-mode. In Insert-mode, an SRH is inserted after the IPv6 header of the BFD packet. In Encaps-mode, the BFD packet is encapsulated in an outer IPv6 header with an SRH.

The naming of these two modes comes from SRv6 Policy headend behaviors, H.Insert [I-D.filsfils-spring-srv6-net-pgm-insertion] and H.Encaps [RFC8986]. The encapsulation of BFD packets for an SRv6 Policy can be viewed as if the headend steers it into the SRv6 Policy.

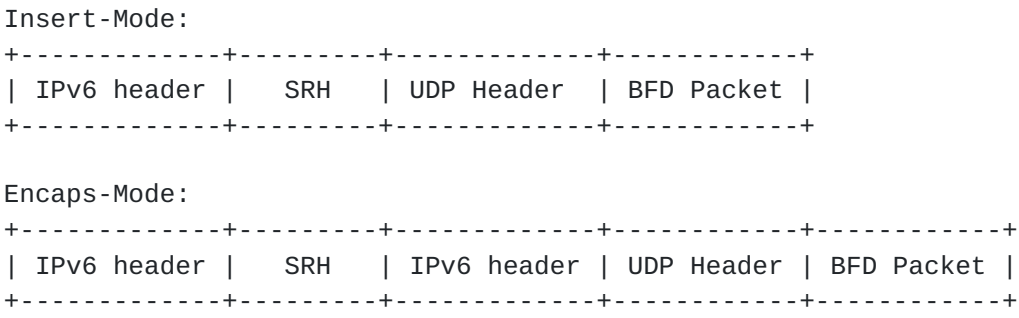

Figure 1: Encapsulation of BFD Packet for SRv6 Policy

How to setup sessions for the segment lists associated with an SRv6 Policy is out of the scope of this document.

2.1. Insert-Mode

In Insert-mode, the encapsulation format of BFD control packet is as follows:

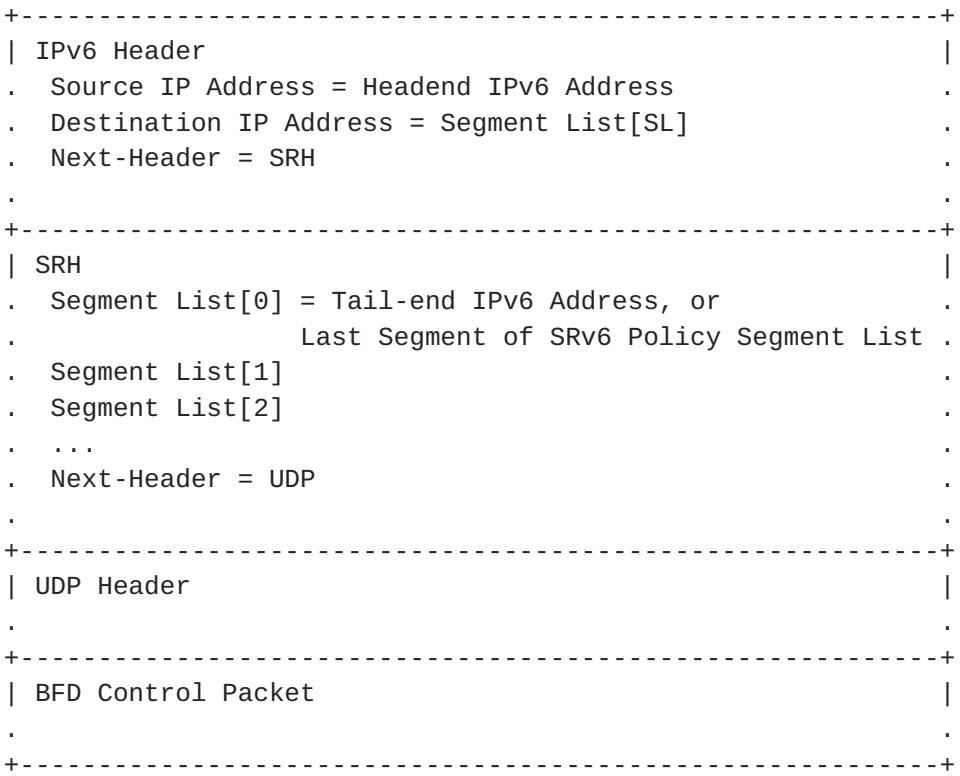

Figure 2: Format of Control Packet in Insert-Mode

In the SRH, the first element of the segment list (Segment List[0]) contains the SRv6 SID or IPv6 address of the tail-end node.

If the last segment of the SRv6 Policy segment list does not belong to the tail-end node, an IPv6 address of tail-end should be added as Segment List[0], while Segment List[1] contains the last segment of the SRv6 Policy segment list. The typical scenarios are as follows:

- The last segment of the SRv6 Policy segment list may be an End.X SID of the penultimate hop. If it is used as Segment List[0], the final destination for the BFD packet is missing.

- The last segment of the SRv6 Policy segment list may be a Binding SID, for example, the application of SRv6 Policy for L3VPN service across multiple domains. If it is used as segment list[0], according to [RFC8986], the node which instantiates the BSID will not perform the encapsulation behavior of the associated SRv6 Policy, but stop processing the SRH and proceed to process the next header in the packet.

Else, the additional tail-end IPv6 address is not necessary, and it can be omitted in order to reduce the SRH size.

2.2. Encaps-Mode

In Encaps-mode, the encapsulation format of BFD control packet is as follows:

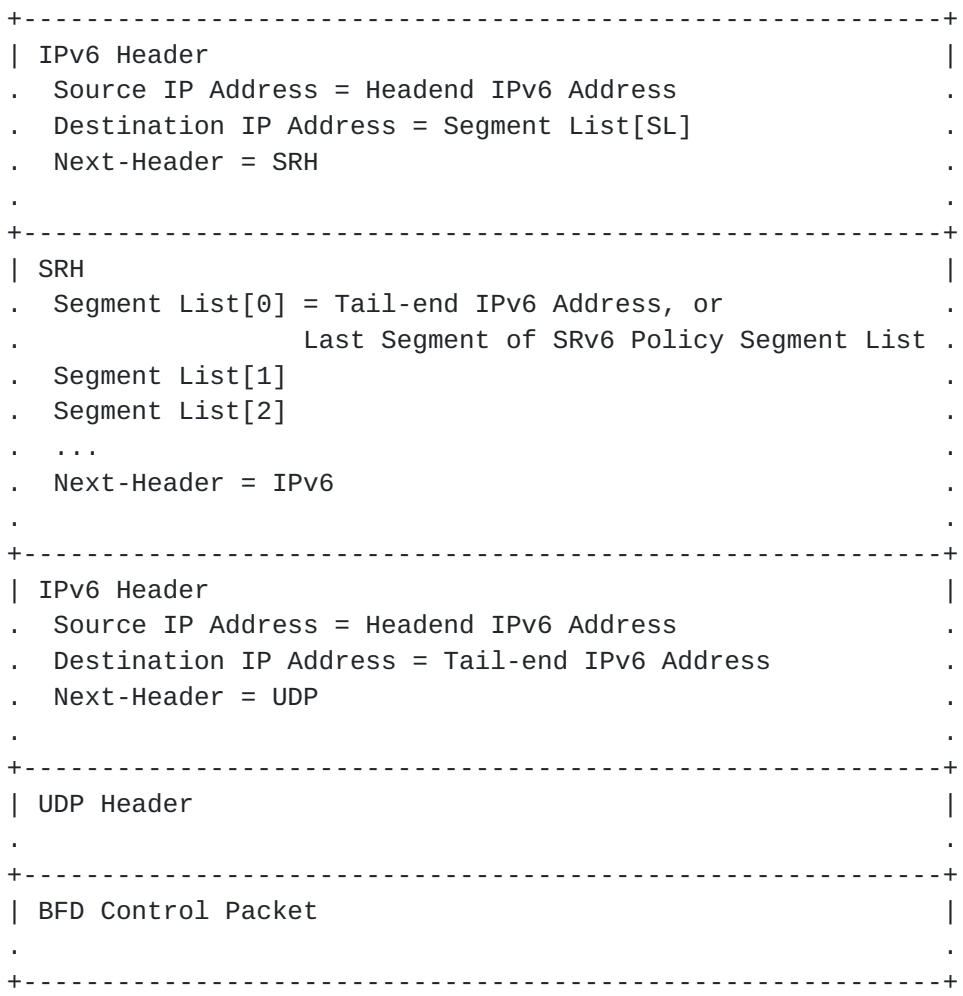

Figure 3: Format of Control Packet in Encaps-Mode

In the SRH, the first element of the segment list (Segment List[0]) contains the SRv6 SID or IPv6 address of the tail-end node.

If the last segment of the SRv6 Policy segment list does not belong to the tail-end node and its function does not include decapsulation of the outer IPv6 header, an IPv6 address of tail-end should be added as Segment List[0], while Segment List[1] contains the last segment of the SRv6 Policy segment list. The typical scenarios are as follows:

- The last segment of the SRv6 Policy may be an End.X SID of the penultimate hop. If it is used as Segment List[0], the penultimate hop needs to remove the outer IPv6 header with all SRH, and forwards the inner IPv6 packet to reflector. If the last segment is with Ultimate Segment Decapsulation (USD) flavor, the penultimate SR endpoint node will perform such decapsulation as defined in [RFC8986]. Otherwise, how to process the packet when the upper-layer header type is IPv6, is not clearly defined in [RFC8986]. It depends on implementation, and may not work well for BFD.

- The last segment of the SRv6 Policy may be a Binding SID, which is the same with the Binding SID case in section 2.1.

Else, the additional tail-end IPv6 address is not necessary, and it can be omitted in order to reduce the SRH size.

3. Choice of Headend and Tail-end IPv6 Addresses

When traffics are steered into an SRv6 Policy, the headend encapsulates the received packets in an outer IPv6 header along with an SRH. The Source Address of the outer IPv6 header is an IPv6 Address of the headend itself which can be routed. It may be a local interface address of the headend used for all SRv6 Policies. Or, different source addresses may be allocated per SRv6 Policy by local configuration.

For the BFD control packet, it is RECOMMENDED to use the headend IPv6 address associated with the SRv6 Policy as the Source Address of (outer) IPv6 header.

An SRv6 Policy is identified through the tuple <headend, color, endpoint>. The endpoint indicates the destination of the policy, and is usually specified as an IPv6 address of the tail-end node.

For the BFD control packet, the headend is RECOMMENDED to choose the endpoint of the SRv6 Policy to be the tail-end IPv6 address which appears in Segment List[0] of SRH or DA of inner IPv6 header, without additional knowledge of the tail-end. In the cases where the endpoint of SRv6 Policy is the unspecified address (:: for IPv6), the tail-end IPv6 Address SHOULD be specified by local configuration or network controller.

4. Checksum in UDP Header

The computation of Checksum in UDP header includes the Destination Address of IPv6 header.

In the encapsulation of Insert-mode, the IPv6 DA may change along the SRv6 forwarding path. When computing the UDP Checksum, the headend should use Segment List[0] in the SRH as the IPv6 DA. It is consistent with the packet received by the final destination, the tail-end node. So, when the final destination processes the UDP header, the verification of Checksum will be passed.

In the encapsulation of Encaps-mode, the computation of UDP Checksum only involves the inner IPv6 header, which does not change en route. No additional action needs to be taken.

5. Control of Adding Tail-end IPv6 Address in SRH

In order to make sure the BFD control packet reaches the tail-end, an implementation MUST add a tail-end IPv6 address as Segment List[0] in the SRH when it is necessary. Otherwise, it MAY be omitted to reduce the SRH size.

Since the headend may not be able to make such a judgment, it is RECOMMENDED that an implementation always adds a tail-end IPv6 address as Segment List[0] in the SRH of the BFD control packet.

6. Example

In the following network, the headend A installs an SRv6 Policy to tail-end D with one segment list <SID-A1, SID-B1, SID-C1>. SID-A1, SID-B1, and SID-C1 are all SRv6 End.X SIDs. Assume that A uses S-BFD to monitor that SRv6 Policy.

Figure 4: example network

The S-BFD control packet in Insert-mode is shown in Figure 5.

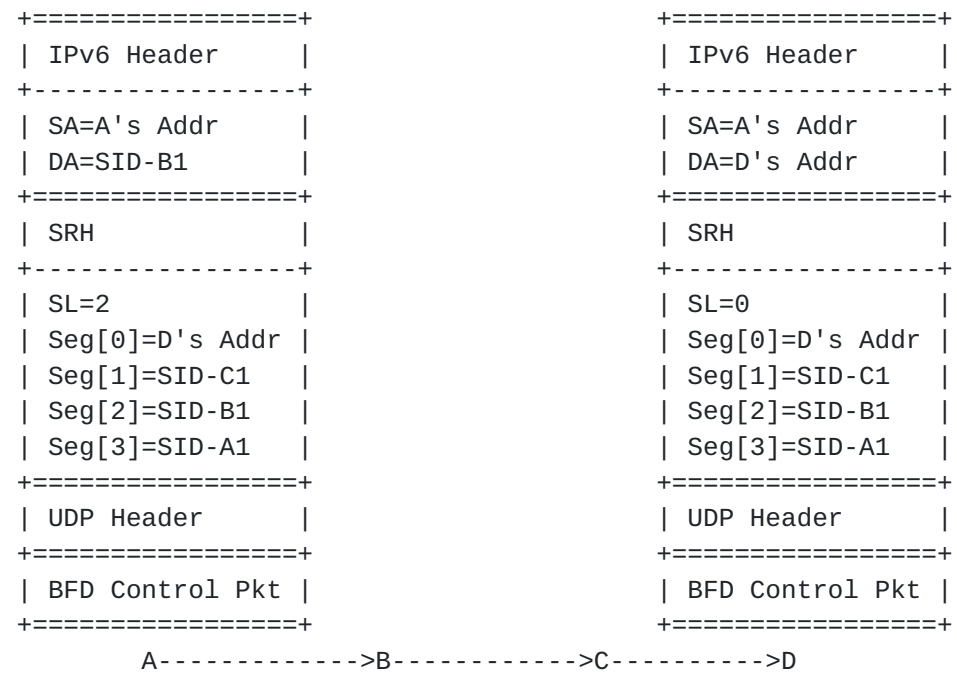

Figure 5: Example of S-BFD Control Packet in Insert-Mode

The S-BFD control packet in Encaps-mode is shown in Figure 6.

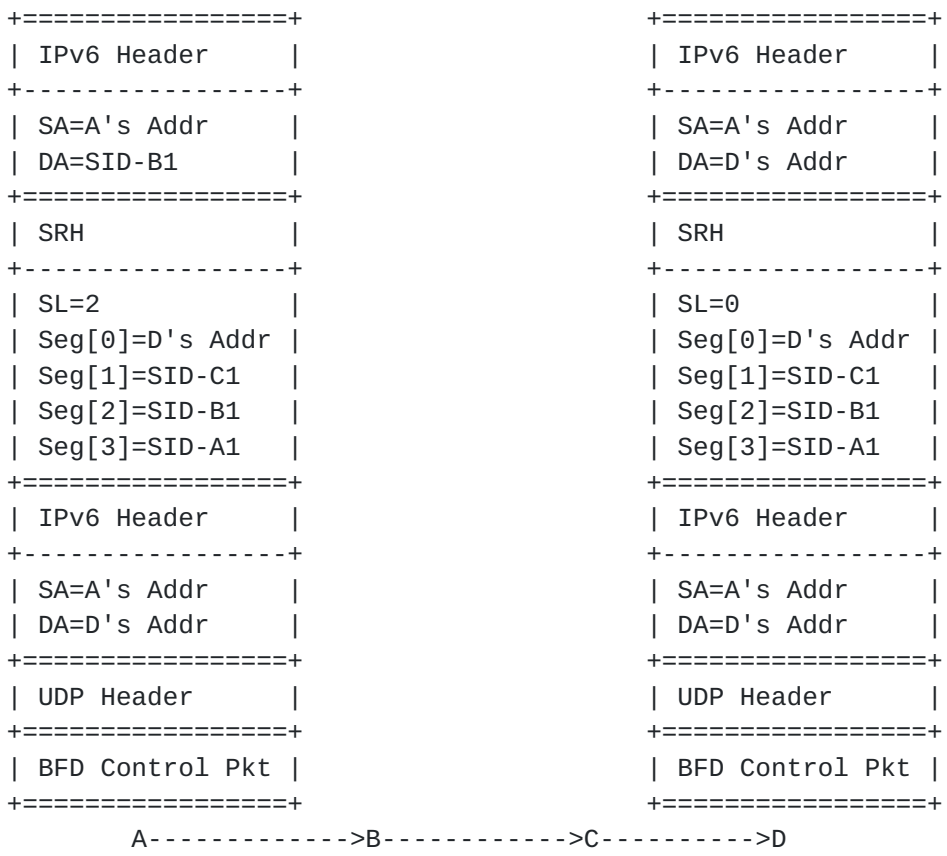

Figure 6: Example of S-BFD Control Packet in Encaps-Mode

7. Security Considerations

TBD.

8. IANA Considerations

This document has no IANA actions.

9. References

9.1. Normative References

[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>.

[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>.

[RFC8402] Filsfils, 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>.

[RFC9256] Filsfils, C., Talaulikar, K., Voyer, D., Bogdanov, A., and P. Mattes, "Segment Routing Policy Architecture", RFC9256, DOI 10.17487/RFC9256, July 2022, <https://datatracker.ietf.org/info/rfc9256>.

[RFC5880] Katz, D. and D. Ward, "Bidirectional Forwarding Detection (BFD)", RFC 5880, DOI 10.17487/RFC5880, June 2010, <http://www.rfc-editor.org/info/rfc5880>.

[RFC7880] Pignataro, C., Ward, D., Akiya, N., Bhatia, M., and S. Pallagatti, "Seamless Bidirectional Forwarding Detection (S-BFD)", RFC 7880, DOI 10.17487/RFC7880, July 2016, <https://www.rfc-editor.org/info/rfc7880>.

[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>.

[RFC8754] Filsfils, C., Ed., Dukes, D., Ed., Previdi, S., Leddy, J., Matsushima, S., and D. Voyer, "IPv6 Segment Routing Header (SRH)", RFC 8754, DOI 10.17487/RFC8754, March 2020, <https://www.rfc-editor.org/info/rfc8754>.

[RFC8986] Filsfils, C., Ed., Camarillo, P., Ed., Leddy, J., Voyer, D., Matsushima, S., and Z. Li, "Segment Routing over IPv6 (SRv6) Network Programming", RFC 8986, DOI 0.17487/RFC8986, February 2021, <https://www.rfc-editor.org/info/rfc8986>.

9.2. Informative References

[I-D.ietf-spring-bfd] Mirsky, G., Tantsura, J., Varlashkin, I., Chen, M., and J. Wenying, " Bidirectional Forwarding Detection (BFD) in Segment Routing Networks Using MPLS Dataplane ", Work in Progress, Internet-Draft, draft-ietf-spring-bfd-08, August 1 2023, <http://www.ietf.org/internet-drafts/draft-ietf-spring-bfd-08.txt>.

[I-D.filsfils-spring-srv6-net-pgm-insertion] Filsfils, C., Camarillo, P., Leddy, J., Voyer, D., Matsushima, S., and Z. Li, "SRv6 NET-PGM extension: Insertion", Work in Progress, Internet-Draft, draft-filsfils-spring-srv6-net-pgm-insertion-09, 16 August 2023, <http://www.ietf.org/internet-drafts/draft-filsfils-spring-srv6-net-pgm-insertion-09.txt>.

[I-D.ietf-pim-sr-p2mp-policy] Voyer, D., Filsfils, C., Parekh, R., Bidgoli, H., and Z. J. Zhang, "Segment Routing Point-to-Multipoint Policy", Work in Progress, Internet-Draft, draft-ietf-pim-sr-p2mp-policy-07, 11 Ocotober 2023, <https://datatracker.ietf.org/doc/html/draft-ietf-pim-sr-p2mp-policy-07>.

Acknowledgements

The authors would like to thank Greg Mirsky for his review and comments of this document.

Authors' Addresses

Yisong Liu
China Mobile
China
Email: liuyisong@chinamobile.com

Weiqiang Cheng
China Mobile
China
Email: chengweiqiang@chinamobile.com

Changwang Lin
New H3C Technologies
China
Email: linchangwang.04414@h3c.com

Mengxiao Chen
New H3C Technologies
China
Email: chen.mengxiao@h3c.com

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