

PIM
Internet-Draft
Intended status: Experimental
Expires: 22 January 2026

T. Eckert, Ed.
Futurewei Technologies USA
Y. Liu
China Mobile
Y. Zhu
China Telecom
C. Lin
New H3C Technologies
21 July 2025

The IPv6 Multicast Routing Headers (MRH)
draft-eckert-pim-mrh-01

Abstract

This document describes the IPv6 extensions to support an experiment in which new IPv6 Routing headers that support stateless replication are implemented and deployed for IPv6 Segment Routing Networks. Collectively, these headers are called Multicast Routing Header (MRH).

One purpose of this experiment is to demonstrate that the MRH can be implemented and deployed in a production network. Another purpose is to encourage replication of the experiment.

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 22 January 2026.

Copyright Notice

Copyright (c) 2025 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 Revised BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Revised BSD License.

Table of Contents

1. Overview	2
2. The Multicast Routing Header (MRH)	3
3. MRH nodes	5
4. MRH sub-type semantics	6
4.1. Destination class MRH sub-type	6
4.2. Steering class MRH sub-type	6
5. MRH Packet processing	6
5.1. On Transit Nodes	6
5.2. On the MRH source node	7
5.3. On Segment endpoint nodes	8
6. MRH sub-types	9
6.1. MRH-bitstring-destination (MRH-bd) Sub-Type format	10
6.2. MRH-steering (MRH-s) Sub-Type format	11
7. References	11
7.1. Normative References	11
7.2. Informative References	12
Appendix A. Longer MRH headers	13
Appendix B. Changelog	14
Authors' Addresses	14

1. Overview

This document introduces IPv6 Routing headers that enable stateless replication of packets in IPv6 or SRv6 networks. These headers are collectively called Multicast Routing Headers (MRH). IPv6 Packets using these Routing headers can be IPv6 multicast packets or IPv6 (unicast) packets. When they are multicast packets, these mechanism can enable to avoid stateful multicast tree building protocols, such as PIM-SM, [RFC7761]. When they are unicast packets that need to be sent to more than one receiver, this mechanism allows to avoid either sender-side replication of these packets or the use of IPv6 multicast to achieve the replication.

The mechanisms in this document are specifically targeted for SRv6 networks where stateless forwarding through ingress node steering of packets via SRH ([RFC8754])) already establishes a network preference to minimize the amount of state required in networks and/or the

desire or need for advanced traffic steering, both of which is also possible for stateless replicated packets through the mechanisms in this memo.

This document intends to support an experiment in which these new IPv6 Routing headers are implemented and deployed for IPv6 Segment Routing Networks or IPv6 network without the desire for other forwarding planes.

One purpose of this experiment is to demonstrate that the MRH can be implemented and deployed in a production network. Another purpose is to encourage replication of the experiment.

This document only specifies the Routing headers and their semantics. The larger context of the experiment is (informationally) described in [I-D.eckert-pim-mrh-experiment].

This document is intended for 6MAN which is authoritative for IPv6 extension headers, but it is initially directed at review in PIM which is responsible for the overall experiment as the Multicast Routing expert group, including Multicast for Segment Routing/SRv6.

2. The Multicast Routing Header (MRH)

The MRH header contains the following fields:

- * Next Header - Defined in [RFC8200].
- * Hdr Ext Len - Defined in [RFC8200].
- * Routing Type - Defined in [RFC8200]. Each MRH sub-type is allocated a new Routing Type by IANA.
- * Segments Left - Defined in [RFC8200].
- * Type-specific Data - Described in [RFC8200].

In the MRH, the Type-specific Data field contains the elements as shown in Figure 1.

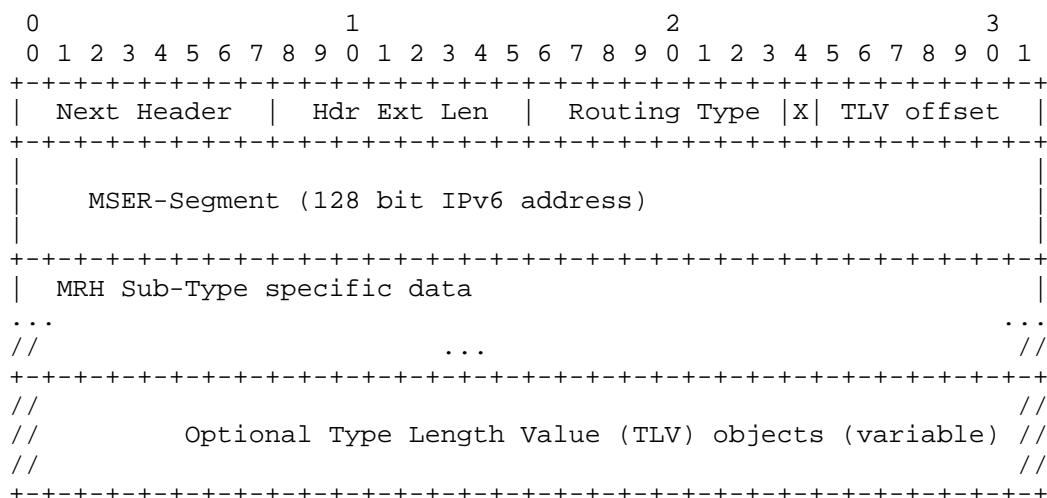


Figure 1: MRH format

The "MSER-Segment" ("Multicast Source Exit Router Segment") carries an IPv6 address. If it is an IPv6 address according to [RFC4291], then the packet is an IPv6 multicast packet.

If the MSER-Segment is not an IPv6 address, then it is a SID according to [RFC8986], and the packet is called a stateless, directed replication, SRv6 (unicast) packet.

MRH Sub-Type specific data contains a compressed encoding of SIDs that permits replication to a set of one or more destination SR nodes. Its encoding is determined by the MRH Sub-Type indicated by the Routing Type field.

The data contained in MSER-Segment does not allow to determine on every Segment endpoint node a calculation of the actual number of segments left, nor would this value be useful. For this reasons, the MRH header field does not include a "Segments Left" field, but that space is instead used to indicate the offset to the optional TLV objects field in 32 bit units. This also makes it necessary for the MRH Sub-Type field to be multiple of 32 bit units in length - or be padded accordingly. If no TLV objects field is present, the value of TLV offset points to the first byte after the end of the MRH header.

Note that the definition of TLV offset makes it never zero, which is necessary so that routers not supporting MRH extension headers will stop processing the packet and raise an ICMPv6 error according to [RFC8200], section 4.4.

The X Flag is for future extensions, see Appendix A. It MUST be set to 0. Packets received with X set to 1 MUST, the node must discard the packet and send an ICMP Parameter Problem, Code 0, message to the packet's Source Address, pointing to the unrecognized Routing Type.

The length of the MRH Sub-Type may change during processing. TLV offset then needs to be updated accordingly. MTU discovery is not used for multicast packets and MRH headers are intended for use in controlled networks, so change in header size is something that can be managed. For example by never sending packets with more than the minimum MTU of 1280 as recommended for multicast packets in [RFC3542] and accordingly setting the MTU in the controlled network so the largest possible headersizes plus 1280 will fit.

The Optional Type Length Value (TLV) objects field is as defined in [RFC8754]. As specified there, its length may change.

Except for the aforementioned fixed value of the Segments left field, routing and forwarding for packets with an MRH routing extension header follows the rules of [RFC8200] as restated and refined by the following definitions.

3. MRH nodes

There are different types of nodes that may be involved in MRH segment routing networks:

- * Source nodes: originate packets with an MRH segment address in the destination address of the IPv6 header
- * Transit nodes: forward packets destined to a remote MRH segment
- * Segment endpoint nodes: process a local segment in the destination address of an IPv6 header.
- * Egress nodes: a subset of the Segment endpoint node that act as IPv6/IPv6-Multicast host to the packet, passing it up to the next header protocol of the packet

Note that SRv6 in [RFC8754] does not explicitly name egress nodes, although every SRH packet has exactly one such Egress node too. This document simply adds that additional terminology for convenience reasons. All other node types have exactly the same specification as in [RFC8754].

4. MRH sub-type semantics

All MRH sub-types operate on the semantic that a packet is forwarded and replicated across a directed tree rooted in the source node of the MRH packet. All edges of the tree are segments, each vertex of the tree is a segment endpoint node. All leaves and a subset of the non-leaf are egress nodes, operating as IPv6 hosts processing the next header protocol of the packet.

4.1. Destination class MRH sub-type

In a "destination" class MRH sub-type, the MRH sub-type specific data encodes only a list of egress node SIDs. All leaves of the tree are egress nodes. Non-egress edges/nodes of the tree are determined by forwarding/replication on segment endpoint nodes. Each segment endpoint node determines next segment endpoint nodes to replicate the packet to in order to reach the subset of egress nodes for the subtree rooted in this node. This is the replication mechanism used for example in [RFC8279].

4.2. Steering class MRH sub-type

In a "steering" class MRH sub-type, the MRH sub-type specific data encodes the complete directed tree with its segment endpoint nodes. Each non-leaf segment endpoint node determines the next segment endpoint nodes to which to replicate the packet to directly from the MRH sub-type specific data without having to examine the encoded tree beyond those next-hops. This is the replication mechanism used for example in [RFC9262].

In a "steering" class MRH sub-type, the non-leaf nodes of the tree are pre-determined by the MRH sub-type data and delivery of the packet to that node can thus be granted. Therefore, the actual property of a node being an egress node does not need to be encoded in the MRH sub-type but can instead be derived from membership of that node in the packets IPv6 multicast and/or future MSER segment SID semantic. This allows for more compact encoding of MRH sub-type data in this case because it eliminates the need to distinguish between a non-leaf segment endpoint node of the tree to be an egress node or not.

5. MRH Packet processing

5.1. On Transit Nodes

Processing of MRH extension headers is as specified in [RFC8754], Section 4.2. with respect to [RFC8200] forwarding of IPv6 packets.

However, additional packet processing in the forwarding plane that operate on IPv6 multicast packets for functions beyond those of IPv6 (unicast) packet forwarding MAY perform the same operation on an IPv6 packet with an MRH extension header which is an IPv6 multicast packet because its MSER-Segment contains an IPv6 multicast group address.

In this case, the IPv6 multicast related function is performed as if the IPv6 multicast group address in the MSER-Segment was present in the IPv6 DA field of the packet, ignoring the IPv6 (unicast) Segment endpoint address in the IPv6 DA field, which is irrelevant to the IPv6 multicast semantic of the packet.

Examples of such functions are statistics gathering of IPv6 multicast packets for IPFIX, attachment of QoS policies to the packet or "ACL" filtering of packet including firewall functions or filtering of scoped IPv6 multicast addresses ([RFC7346]).

5.2. On the MRH source node

- * A source node steers a packet into an MRH/SR policy that translates into one or more of the following set of data elements, each one required for one copy of the packet.
 - An MSR sub-type header representing in a compressed form a set of Egress node SIDs (destination sub-type) or a tree of segment endpoint node SIDs with optional Egress nodes (steering sub-type)
 - An IPv6 source address belonging to the source node, which may be a SID, or which may need to be associated with semantics of the MSR sub-type header or the packet. For example, for an IPv6 multicast SSM packet as in[RFC3306], [RFC3569], the SA needs to be set according to the desired (S,G) channel.
 - An IPv6 destination address, which may be an IPv6 multicast group address (G) or a SID to be processed by all Egress nodes.
 - TLV (including HMAC).
 - Optionally an IPv6 destination address for the first Segment endpoint node.

The MRH/SR policy needs more than one data element sets and hence packet copies when the desired list of egress nodes and/or representation of the tree of segment endpoint does not fit into a single MRH header.

- * The source node sets the IPv6 header and MRH extension header as follows
 - The SA of the packet is set to the IPv6 source address
 - The MSER-segment is set to the IPv6 destination address.
 - The Next Header and Hdr Ext Len fields are set as specified in [RFC8200]. The Routing Type field is set to the value assigned for the MSR sub-type.
 - The MSR sub-type header is inserted into its field. The Segments left field is set to the length of the MSR sub-type header plus 32 bits in units of 32 bits.
 - If the SR policy includes a first Segment endpoint node IPv6 address, then the packets DA is set to that address, and the packet is sent to it.
 - If the SR policy does not include such an IPv6 address, the packets DA is logically left unset and the packet is passed to the nodes MRH Segment endpoint processing steps.

5.3. On Segment endpoint nodes

This document requires/uses only a single SRv6 SID for the reception of IPv6 packets with MRH type routing extension header. This SID can also be used for other purposes if the node can distinguish the semantic based on the routing header it contains - aka: for packets without an MRH type routing extension header.

- * If local processing requires TLV processing, perform TLV processing as described in in [RFC8754].
- * Determine if this node is an egress node for the packet as follows.
 - If the MRH sub-type data explicitly encodes this node as an egress node.
 - If the node has membership for the IPv6 address in the MSER segment field. This can be the case when it is an IPv6 ASM multicast group that this node has joined to, or if the node has subscribed to the IP6 SSM channel indicated by (SA,MSER-segment), or if the MSER-segment indicates a non multicast SID that the node is subscribed to. The rules for subscription to non multicast SID are outside the scope of this document.

- * If the node is egress node for the packet, create a copy of the packet for which to proceed to process the next header in the packet, whose type is identified by the Next Header field in the routing header.
- * Reduce the Hop-Count of the packet by 1.
- * If the hop count is ≤ 1 , determine from the MSER segment SID whether ICMPv6 processing is desired.
 - If the MSER segment SID is IPv6 multicast, no ICMPv6 processing is desired.
 - Otherwise, the desire for ICMPv6 processing is defined by the SID.
 - If ICMPv6 processing is desired, Send an ICMP Time Exceeded -- Hop Limit Exceeded in Transit message to the Source Address
 - discard the packet, stop processing
- * Determine from the MRH sub-type data a list of 0 or more next-hop segments of the tree and perform the following processing for each of those next-hop segments
 - Create a copy of the packet
 - Change the DA of the packet to the Segment Endpoint SID for that next-hop node.
 - Rewrite the MRH sub-type data according to the processing rule of that MRH sub-type. Typically, in a destination MRH sub-type this consists in eliminating from the encoded Egress nodes set those that are not reachable via this next-hop segment, and for a steering MRH sub-type in adjusting offset pointers in the MRH sub-type data to point to the subtree rooted in the next-hop node.
- * Discard packet, terminate processing

6. MRH sub-types

6.1. MRH-bitstring-destination (MRH-bd) Sub-Type format

MRH-bd uses the packet forwarding data structures (BIFT) and forwarding rules of [RFC8279]. IPv6 packets with MRH-sd extension header are an [RFC8279] compliant encapsulation to support BIER architecture ([RFC8279] packet forwarding, and it uses all forwarding rules unmodified. [RFC8296] encapsulation is not used.

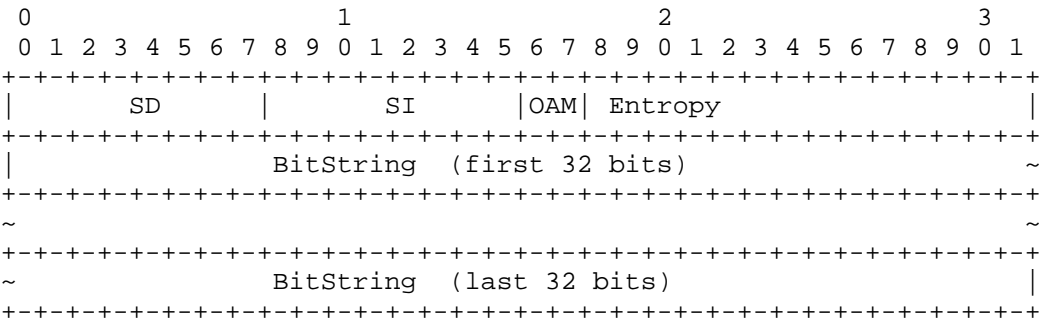


Figure 2: MRH-bd Sub-Type format

Figure 2 shows the field of the MRH Sub-type specific data for the MRH

- * SD is the sub-domain as specified in [RFC8279]
- * SI is the Set Identifier as specified in [RFC8279]
- * Entropy is as specified in [RFC8279].
- * BitString is as specified in [RFC8279].
- * OAM is for future use with OAM procedures.

The length of the BitString MUST be a multiple of 32 bits. It SHOULD be one of the following value for compatibility with potentially pre-existing {RFC8279}} forwarding hardware:

64, 128, 256, 512, 1024, (2048, 4096)

Nodes supporting MRH-sd MUST support a BSL of 256 bits. Note that the length of 2048 and 4096 are not possible in an MRH extension header due to the 255 byte size limit of IPv6 extension headers. See Appendix A.

Population of the [RFC8279], section 6.4 forwarding tables (BIFT) is outside the scope of this specification. For the purpose of performing the rewriting of the IPv6 DA on Segment endpoint nodes, the BFR-NBR in the BIFT needs to be the MRH-sd SID of the next-hop MRH-sd node.

6.2. MRH-steering (MRH-s) Sub-Type format

This sub-type is TBD. There are currently multiple competing proposals considered, including RTS ([I-D.eckert-pim-rtts-forwarding] and [I-D.liu-pim-msr6-encapsulation-and-forwarding]).

It should be noted, that [RFC9262] will explicitly not be considered, because that work was focussed on re-using the [RFC8279] forwarding plane as much as possible, and this lead to a lot of management and scalability issues which would make experimentation not with it not very useful. [RFC8279] can work fine in smaller scale deployments (networks and/or size of trees), but the purpose of this experiment is also to find better solutions for future hardware forwarding engines.

7. References

7.1. Normative References

- [RFC3306] Haberman, B. and D. Thaler, "Unicast-Prefix-based IPv6 Multicast Addresses", RFC 3306, DOI 10.17487/RFC3306, August 2002, <<https://www.rfc-editor.org/rfc/rfc3306>>.
- [RFC4291] Hinden, R. and S. Deering, "IP Version 6 Addressing Architecture", RFC 4291, DOI 10.17487/RFC4291, February 2006, <<https://www.rfc-editor.org/rfc/rfc4291>>.
- [RFC7346] Droms, R., "IPv6 Multicast Address Scopes", RFC 7346, DOI 10.17487/RFC7346, August 2014, <<https://www.rfc-editor.org/rfc/rfc7346>>.
- [RFC8200] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", STD 86, RFC 8200, DOI 10.17487/RFC8200, July 2017, <<https://www.rfc-editor.org/rfc/rfc8200>>.
- [RFC8279] Wijnands, IJ., Ed., Rosen, E., Ed., Dolganow, A., Przygienda, T., and S. Aldrin, "Multicast Using Bit Index Explicit Replication (BIER)", RFC 8279, DOI 10.17487/RFC8279, November 2017, <<https://www.rfc-editor.org/rfc/rfc8279>>.

- [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/rfc/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 10.17487/RFC8986, February 2021, <<https://www.rfc-editor.org/rfc/rfc8986>>.

7.2. Informative References

- [I-D.eckert-pim-mrh-experiment]
Eckert, T. T. and Y. Liu, "Proposal for experimentation with stateless IPv6 multicast source routing in IPv6-only SRv6 networks via new IPv6 Multicast Routing Headers (MRH)", Work in Progress, Internet-Draft, draft-eckert-pim-mrh-experiment-01, 7 July 2025, <<https://datatracker.ietf.org/doc/html/draft-eckert-pim-mrh-experiment-01>>.
- [I-D.eckert-pim-rtts-forwarding]
Eckert, T. T., Menth, M., Lindner, S., and Y. Liu, "Stateless Multicast Replication with Segment Routed Recursive Tree Structures (RTS)", Work in Progress, Internet-Draft, draft-eckert-pim-rtts-forwarding-03, 6 July 2025, <<https://datatracker.ietf.org/doc/html/draft-eckert-pim-rtts-forwarding-03>>.
- [I-D.liu-pim-msr6-encapsulation-and-forwarding]
Liu, Y., Geng, X., and C. Lin, "Encapsulation and Forwarding Process for IPv6 Multicast Source Routing (MSR6) Data Plane", Work in Progress, Internet-Draft, draft-liu-pim-msr6-encapsulation-and-forwarding-01, 24 April 2025, <<https://datatracker.ietf.org/doc/html/draft-liu-pim-msr6-encapsulation-and-forwarding-01>>.
- [RFC3542] Stevens, W., Thomas, M., Nordmark, E., and T. Jinmei, "Advanced Sockets Application Program Interface (API) for IPv6", RFC 3542, DOI 10.17487/RFC3542, May 2003, <<https://www.rfc-editor.org/rfc/rfc3542>>.
- [RFC3569] Bhattacharyya, S., Ed., "An Overview of Source-Specific Multicast (SSM)", RFC 3569, DOI 10.17487/RFC3569, July 2003, <<https://www.rfc-editor.org/rfc/rfc3569>>.

- [RFC7761] Fenner, B., Handley, M., Holbrook, H., Kouvelas, I., Parekh, R., Zhang, Z., and L. Zheng, "Protocol Independent Multicast - Sparse Mode (PIM-SM): Protocol Specification (Revised)", STD 83, RFC 7761, DOI 10.17487/RFC7761, March 2016, <<https://www.rfc-editor.org/rfc/rfc7761>>.
- [RFC8296] Wijnands, IJ., Ed., Rosen, E., Ed., Dolganow, A., Tantsura, J., Aldrin, S., and I. Meilik, "Encapsulation for Bit Index Explicit Replication (BIER) in MPLS and Non-MPLS Networks", RFC 8296, DOI 10.17487/RFC8296, January 2018, <<https://www.rfc-editor.org/rfc/rfc8296>>.
- [RFC9262] Eckert, T., Ed., Menth, M., and G. Cauchie, "Tree Engineering for Bit Index Explicit Replication (BIER-TE)", RFC 9262, DOI 10.17487/RFC9262, October 2022, <<https://www.rfc-editor.org/rfc/rfc9262>>.

Appendix A. Longer MRH headers

Currently, IPv6 extension headers can only be up to 255 byte in size. It is also perceived common understanding that high-speed IPv6 routers, which are also targeted by this memo can not well support lookup into packets of more than 512 bytes, and that examining a long header may reduce packet throughput performance. For example, an [RFC8279] Bitstring has to be completely examined by a router, so longer Bitstrings run into this problem.

However, with segment tree encoding in MRH sub-types, a single Segment endpoint nodes does not need to examine the whole header anymore, but only a consecutive smaller part of a potentially much larger header. Likewise, if each node removes its part of the MRH sub-type data for the copy sent to the next-hop node, then each node will also have its part of the header start near the start of MRH header, thus no look deep into the packet is necessary. Likewise, router forwarding engines become like general purpose CPU, the less problematic larger offsets into packet memory will be an issue, but only the total amount of processing needed to be done on the memory accessed.

Finally, is it worth to have a larger header instead of sending multiple packets ? Yes, for multicast packets it is. Consider that a 1024 byte long extension header could address for example 4 times as many destinations as a 256 byte extension header, so the bandwidth use near the source of the packet is significantly reduced by as much as a factor of 4.

One possible option would be to use the X field in the MRH header to indicate that the Hdr Ext Len indicates the length of the extension header in 32 or 64 bits.

Appendix B. Changelog

01: Added new co-authors

00: initial version.

Authors' Addresses

Toerless Eckert (editor)
Futurewei Technologies USA
United States of America
Email: tte@cs.fau.de

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

Yongqing Zhu
China Telecom
China
Email: zhuyq8@chinatelecom.cn

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