

Workgroup: MPLS Working Group
Internet-Draft: draft-gandhi-mpls-ioam-05
Published: 3 October 2022
Intended Status: Standards Track
Expires: 6 April 2023

A R. Gandhi, Ed. Z. Ali
uCisco Systems, Inc. Cisco Systems, Inc.
t
h
o
r
s
:
F. Brockners B. Wen B. Decraene
Cisco Systems, Inc. Comcast Orange
H. Song V. Kozak
Futurewei Technologies Comcast

MPLS Data Plane Encapsulation for In-situ OAM Data

Abstract

In-situ Operations, Administration, and Maintenance (IOAM) is used for recording and collecting operational and telemetry information while the packet traverses a path between two points in the network. This document defines how IOAM data fields are transported with MPLS data plane encapsulation using MPLS Network Action (MNA).

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 6 April 2023.

Copyright Notice

Copyright (c) 2022 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. Introduction](#)
- [2. Conventions](#)
 - [2.1. Requirement Language](#)
 - [2.2. Abbreviations](#)
- [3. MPLS Extensions for IOAM Data Fields](#)
 - [3.1. IOAM Extension Header](#)
 - [3.2. Post-Stack Network Action Presence Indicators](#)
- [4. Edge-to-Edge IOAM Network Action](#)
 - [4.1. Post-Stack Network Action Presence Indicator](#)
 - [4.2. Procedure for Edge-to-Edge IOAM Network Action](#)
- [5. Hop-By-Hop IOAM Network Action](#)
 - [5.1. Hop-By-Hop Post-Stack Network Action Processing Indicator](#)
 - [5.2. Procedure for Hop-By-Hop IOAM Network Action](#)
 - [5.3. Hop-By-Hop and Edge-To-Edge IOAM Extension Headers](#)
- [6. Considerations for IOAM](#)
 - [6.1. Considerations for ECMP](#)
 - [6.2. Node Capability](#)
 - [6.3. Nested MPLS Encapsulation](#)
- [7. Security Considerations](#)
- [8. IANA Considerations](#)
- [9. References](#)
 - [9.1. Normative References](#)
 - [9.2. Informative References](#)
- [Acknowledgements](#)
- [Authors' Addresses](#)

1. Introduction

In-situ Operations, Administration, and Maintenance (IOAM) is used for recording and collecting operational and telemetry information while the packet traverses a path between two points in the network. The term "in-situ" refers to the fact that the IOAM data fields are added to the data packets rather than being sent within the probe packets specifically dedicated to OAM. The IOAM data fields are defined in [RFC9197]. The IOAM data fields are further updated in [I-D.ietf-ippm-ioam-direct-export] for direct export use-cases.

This document defines how IOAM data fields are transported with MPLS data plane encapsulations using MPLS Network Action (MNA) [I-D.ietf-mpls-mna-fwk].

Note: This document uses MNA encoding defined in Individual Internet Drafts as examples. The MNA encoding will be updated to align with the MNA solutions in the WG adopted Internet Drafts.

2. Conventions

2.1. Requirement Language

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 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

2.2. Abbreviations

The MPLS Network Action (MNA) terminology defined in [[I-D.ietf-mpls-mna-fwk](#)] are used in this document.

Abbreviations used in this document:

PNI Post-Stack Network Action Presence Indicator

ECMP Equal Cost Multi-Path

E2E Edge-To-Edge

HBH Hop-By-Hop

HPI Hop-By-Hop Post-Stack Network Action Processing Indicator

IOAM In-situ Operations, Administration, and Maintenance

MPLS Multiprotocol Label Switching

MNA MPLS Network Action

OAM Operations, Administration, and Maintenance

POT Proof-of-Transit

3. MPLS Extensions for IOAM Data Fields

3.1. IOAM Extension Header

The IOAM extension headers containing different IOAM-Data-Fields are added in the MPLS packet as shown in Figure 1. The IOAM-Data-Fields MUST follow the definitions corresponding to the IOAM-Option-Types (e.g., see Section 4.4 of [[RFC9197](#)] and Section 3 of [[I-D.ietf-ippm-ioam-direct-export](#)]).

An IOAM extension header uses Extension Header (EH) defined in [[I-D.song-mpls-extension-header](#)].

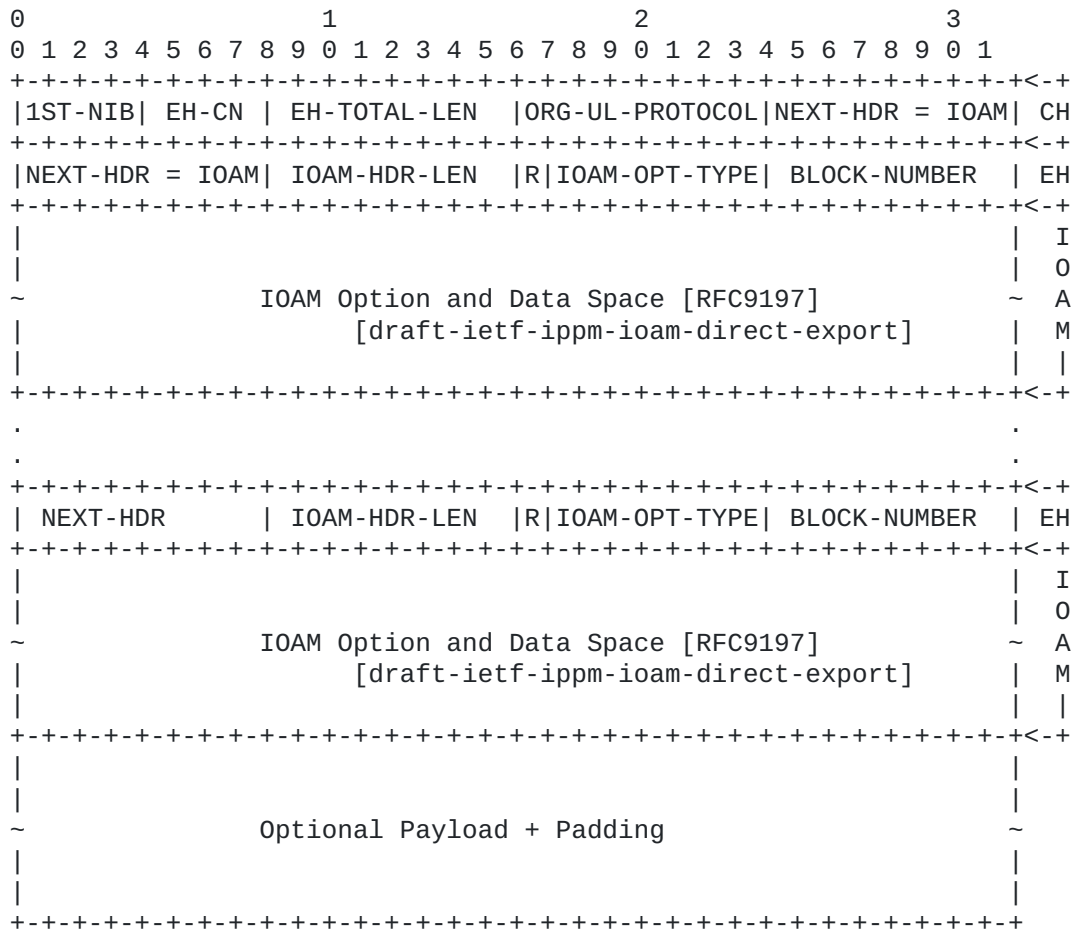


Figure 1: Extension Headers with IOAM Data Fields

The 4-Octet Common Header (CH) is added with the following fields in the MPLS header as defined in [[I-D.song-mpls-extension-header](#)]:

1ST-NIBBLE (R): The first 4-bit is Reserved (value TBA4).

EH-CN (EHC): The EH Count, number of EHs carried in the packet.

EH-TOTAL-LEN (EHTL): Total Length of EHs in the packet in 4-octet units.

ORIG-UL-PROTOCOL (OUL): The Original Upper Layer Protocol.

NEXT-HDR (NH): The Next Header protocol. For IOAM as Next Header, value TBA2 is to be assigned by IANA.

IOAM-Data-Fields are added in EHs containing the following fields:

IOAM-OPT-TYPE:

8-bit field defining the IOAM Option type, as defined in the "IOAM Option-Type Registry" specified in [[RFC9197](#)] and [[I-D.ietf-ippm-ioam-direct-export](#)]).

IOAM-HDR-LEN: 8-bit unsigned integer. Length of the IOAM Header in 4-octet units.

IOAM Option and Data Space: IOAM-Data-Fields as specified by the IOAM-OPT-Type field. IOAM-Data-Fields are defined corresponding to the IOAM-Option-Type (e.g., see Section 4.4 of [[RFC9197](#)] and Section 3 of [[I-D.ietf-ippm-ioam-direct-export](#)]).

BLOCK-NUMBER: The Block Number can be used to aggregate the IOAM data collected in data plane, e.g., to compute measurement metrics for each block of a data flow. It is also used to correlate the IOAM data on different nodes.

3.2. Post-Stack Network Action Presence Indicators

A Post-Stack Network Action Presence Indicator MUST be added to indicate the presence of the IOAM-Data-Fields in the MPLS header. If both edge and intermediate nodes need to process IOAM data then both the Post-Stack Network Action Presence Indicator and the HBH Network Action Processing Indicator MUST be added. The HBH Network Action Processing Indicator allows to optimize the IOAM processing on the intermediate nodes and avoids the need to parse all IOAM-Data-Fields.

A flag called PNI (Post-Stack Network Action Presence Indicator) in the TTL field defined in [[I-D.jags-mpls-mna-hdr](#)] is used in this document to indicate the presence of IOAM Post-Stack Network Action and Ancillary Data.

A flag called HPI (Hop-By-Hop Post-Stack Network Action Processing Indicator) in the TTL field defined in [[I-D.jags-mpls-mna-hdr](#)] is used to indicate that HBH processing is required for the Post-Stack Network Action and the associated Ancillary Data.

The MNA Label used in this document is a Base Special Purpose Label (bSPL value TBA1 to be assigned by IANA), and is defined as Network Action Sub-Stack Indicator (NASI) in [[I-D.jags-mpls-mna-hdr](#)].

4. Edge-to-Edge IOAM Network Action

4.1. Post-Stack Network Action Presence Indicator

The Post-Stack Network Action Presence Indicator is set to "1" to indicate the presence of the IOAM-Data-Fields in the MPLS header as shown in Figure 2.

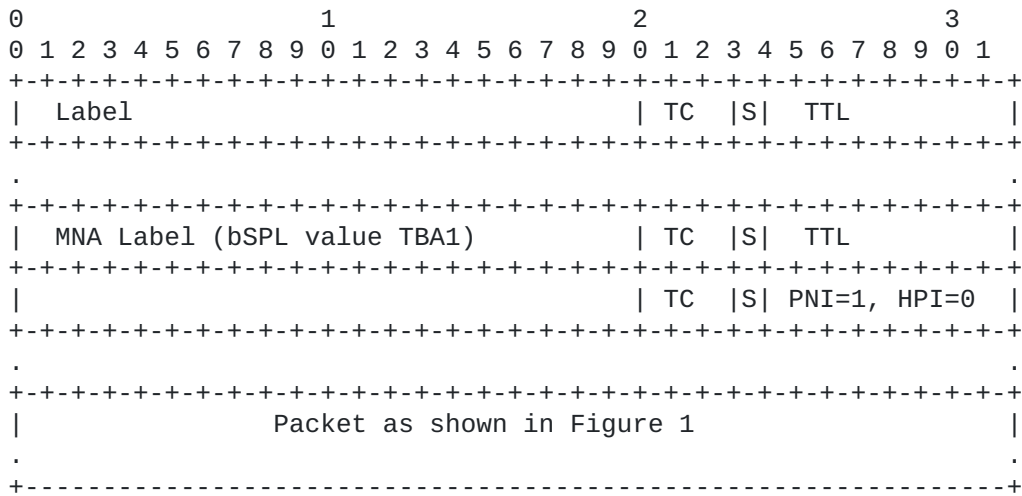


Figure 2: Example MPLS Encapsulation for E2E IOAM

The E2E IOAM-Data-Fields carry the Option-Type(s) that require processing on the encapsulating and decapsulating nodes only. The IOAM Option-Type carried can be IOAM Edge-to-Edge Option-Type (value 3) [[RFC9197](#)] as well as Direct Export (DEX) Option-Type (value 4, to be assigned by IANA) defined in [[I-D.ietf-ippm-ioam-direct-export](#)]. The E2E IOAM-Data-Fields SHOULD NOT carry any IOAM Option-Type that require IOAM processing on the intermediate nodes as it will not be processed by them since HPI flag is not set.

4.2. Procedure for Edge-to-Edge IOAM Network Action

The E2E IOM Network Action procedure is summarized as following:

- *The encapsulating node inserts the MNA Label (bSPL value TBA1) with the Post-Stack Network Action Presence Indicator (TTL Flag PNI) set to "1" below the label whose FEC is the decapsulating node and one or more IOAM-Data-Fields in the MPLS packet.
- *The intermediate nodes do not process IOAM-Data-Fields.
- *The penultimate node MUST NOT remove the MNA Sub-stack from the MPLS header so that the MNA Sub-stack is received at the decapsulating node. This encapsulating node adds required MPLS header so that the received packet on the penultimate node does not have MNA Sub-stack at the top of the Label stack.
- *The decapsulating node MAY punt a copy of the packet with the receive timestamp to the slow path for IOAM-Data-Fields processing when the node recognizes the PNI is set to "1". The receive timestamp is required by the various E2E OAM use-cases, including streaming telemetry. Note that the packet is not necessarily punted to the control-plane.
- *The decapsulating node processes the IOAM-Data-Fields using the procedures defined in [[RFC9197](#)]. An example of IOAM processing is to export the IOAM-Data-Fields, send IOAM-Data-Fields via streaming telemetry, etc.

*The decapsulating node MUST remove the IOAM-Data-Fields from the received packet. The decapsulated packet is forwarded downstream or terminated locally similar to the regular IOAM-Data-Fields.

5. Hop-By-Hop IOAM Network Action

5.1. Hop-By-Hop Post-Stack Network Action Processing Indicator

The Post-Stack Network Action Presence Indicator (TTL Flag PNI) along with Hop-By-Hop Processing Indicator (TTL Flag HPI) are set to "1" to indicate the presence of the HBH IOAM-Data-Fields in the MPLS header as shown in Figure 3.

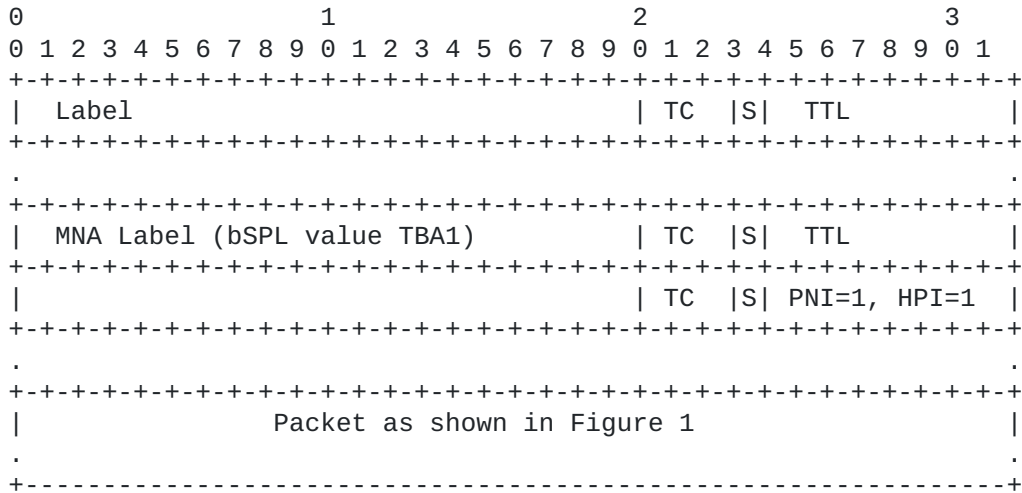


Figure 3: Example MPLS Encapsulation for HBH IOAM

The HBH IOAM-Data-Fields carry the Option-Type(s) that require processing at the intermediate and/or encapsulating and decapsulating nodes. The IOAM Option-Type carried can be IOAM Pre-allocated Trace Option-Type (value 0), IOAM Incremental Trace Option-Type (value 1) and IOAM Proof of Transit (POT) Option-Type (value 2), and Edge-to-Edge Option-Type (value 3) defined in [RFC9197] as well as Direct Export (DEX) Option-Type (value 4), to be assigned by IANA) defined in [I-D.ietf-ippm-ioam-direct-export].

5.2. Procedure for Hop-By-Hop IOAM Network Action

The Hop-By-Hop IOAM Network Action procedure is summarized as following:

*The encapsulating node inserts the MNA Label (bSPL value TBA1) with the Post-Stack Network Action Presence Indicator (TTL Flag PNI) and HBH Processing Indicator (TTL Flag HPI) set to "1" below the label whose FEC is the decapsulating node and one or more IOAM-Data-Fields in the MPLS packet.

*The intermediate node enabled with HBH IOAM function processes the data packet including the IOAM-Data-Fields as defined in [RFC9197] and [I-D.ietf-ippm-ioam-direct-export] when the node recognizes the HBH Processing Indicator in the MPLS header.

*The intermediate node MAY punt a copy of the packet with the receive timestamp to the slow path for IOAM-Data-Fields processing when the node recognizes the HBH Processing Indicator. The receive timestamp is required by the various HBH OAM use-cases, including streaming telemetry. Note that the packet is not necessarily punted to the control-plane.

*The intermediate node forwards a copy of the processed data packet downstream.

*The behavior on the penultimate node is same as E2E case.

*The processing on the decapsulating node is same as E2E case.

5.3. Hop-By-Hop and Edge-To-Edge IOAM Extension Headers

Both HBH and E2E IOAM EHs may be added in an MPLS header. In this case the HBH IOAM EHs MUST be added after the BOS and before the E2E IOAM EHs. When an intermediate node encounters the first E2E IOAM EH, it will stop processing of the E2E EH.

6. Considerations for IOAM

6.1. Considerations for ECMP

The encapsulating node needs to make sure the IOAM-Data-Fields do not start with a well-known IP Version Number (e.g. 0x4 for IPv4 and 0x6 for IPv6) as that can alter the hashing function for ECMP that uses the IP header. This is achieved by using the Common Header (CH) with a different IP Version Number (value TBA4) after the MPLS label stack.

6.2. Node Capability

The decapsulating node that has to remove the IOAM-Data-Fields and perform the IOAM function may not be capable of supporting it. The encapsulating node needs to know if the decapsulating node can support the IOAM function. The signaling extension for this capability exchange is outside the scope of this document.

The intermediate node that is not capable of supporting the IOAM functions defined in this document, can simply skip the IOAM processing.

The node that does not recognize the MNA Label received at the top of the label stack will drop the packet.

6.3. Nested MPLS Encapsulation

When a packet is received with MPLS Encapsulated IOAM Network Action, the nested MPLS encapsulating node that supports a different IOAM Network Action, the node MUST add a new MNA Label (with Sub-Stack) with the supported IOAM Network Action as part of the new MPLS encapsulation.

7. Security Considerations

The security considerations of IOAM in general are discussed in [RFC9197] and apply to the procedure defined in this document.

IOAM is considered a "per domain" feature, where one or several operators decide on configuring IOAM according to their needs. IOAM is intended for deployment in limited domains [RFC8799]. As such, it assumes that a node involved in IOAM operation has previously verified the integrity of the path. Still, operators need to properly secure the IOAM domain to avoid malicious configuration and use, which could include injecting malicious IOAM packets into the domain.

8. IANA Considerations

The encoding of Next-Header uses the same value registry for IPv4/IPv6 protocol numbers. The value TBA2 for IOAM as Next Header shall be assigned by IANA from the "Protocol Numbers" registry (<https://www.iana.org/assignments/protocol-numbers/protocol-numbers.xhtml>).

9. References

9.1. Normative References

[I-D.ietf-ippm-ioam-direct-export]

Song, H., Gafni, B., Brockners, F., Bhandari, S., and T. Mizrahi, "In-situ OAM Direct Exporting", Work in Progress, Internet-Draft, draft-ietf-ippm-ioam-direct-export-11, 23 September 2022, <<https://www.ietf.org/archive/id/draft-ietf-ippm-ioam-direct-export-11.txt>>.

[I-D.jags-mpls-mna-hdr]

Rajamanickam, J., Gandhi, R., Bhattacharya, J., Decraene, B., Zigler, R., Cheng, W., Jalil, L., and H. Song, "MPLS Network Action Header Encodings", Work in Progress, Internet-Draft, draft-jags-mpls-mna-hdr-01, 25 July 2022, <<https://www.ietf.org/archive/id/draft-jags-mpls-mna-hdr-01.txt>>.

[I-D.song-mpls-extension-header]

Song, H., Li, Z., Zhou, T., Andersson, L., Zhang, Z., Gandhi, R., Rajamanickam, J., and J. Bhattacharya, "MPLS Extension Header", Work in Progress, Internet-Draft, draft-song-mpls-extension-header-10, 10 September 2022, <<https://www.ietf.org/archive/id/draft-song-mpls-extension-header-10.txt>>.

[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>>.

[RFC9197] Brockners, F., Bhandari, S., and T. Mizrahi, "Data Fields for In Situ Operations, Administration, and Maintenance", RFC 9197, May 2022, <<https://www.rfc-editor.org/info/rfc9197>>.

9.2. Informative References

[I-D.ietf-mpls-mna-fwk] Andersson, L., Bryant, S., Bocci, M., and T. Li, "MPLS Network Actions Framework", Work in Progress, Internet-Draft, draft-ietf-mpls-mna-fwk-01.txt, 8 September 2022, <<https://www.ietf.org/archive/id/draft-ietf-mpls-mna-fwk-01.txt>>.

[RFC8799] Carpenter, B. and B. Liu, "Limited Domains and Internet Protocols", RFC 8799, DOI 10.17487/RFC8799, July 2020, <<https://www.rfc-editor.org/info/rfc8799>>.

Acknowledgements

The authors would like to thank Patrick Khordoc, Sagar Soni, Shwetha Bhandari, Clarence Filsfils, and Vengada Prasad Govindan for the discussions on IOAM. The authors would also like to thank Tarek Saad, Loa Andersson, Greg Mirsky, Stewart Bryant, Xiao Min, and Cheng Li for providing many useful comments. The authors would also like to thank Mach Chen, Andrew Malis, Matthew Bocci, and Nick Delregno for the MPLS-RT reviews.

Authors' Addresses

Rakesh Gandhi (editor)
Cisco Systems, Inc.
Canada

Email: rgandhi@cisco.com

Zafar Ali
Cisco Systems, Inc.

Email: zali@cisco.com

Frank Brockners
Cisco Systems, Inc.
Hansaallee 249, 3rd Floor
DUESSELDORF, NORDRHEIN-WESTFALEN 40549
Germany

Email: fbrockne@cisco.com

Bin Wen
Comcast

Email: Bin_Wen@cable.comcast.com

Bruno Decraene
Orange

Email: bruno.decraene@orange.com

Haoyu Song
Futurewei Technologies
United States of America

Email: haoyu.song@futurewei.com

Voitek Kozak
Comcast

Email: Voitek_Kozak@comcast.com