TEAS Working Group
Internet Draft
Intended status: Standards Track
Expires: September 03, 2025

L. Gong
China Mobile
C. Lin
New H3C Technologies

March 03, 2025

Operations, Administration and Maintenance (OAM) for Network Resource Partition (NRP) in SR

draft-gong-teas-spring-nrp-oam-00

Abstract

A Network Resource Partition (NRP) represents a subset of network resources and associated policies within the underlay network.

This document describes the implementation of the Operations, Administration, and Maintenance (OAM) mechanism for NRPs in Segment Routing (SR) networks. By extending existing OAM mechanisms such as ping, traceroute, Bidirectional Forwarding Detection (BFD), and Simple Two-way Active Measurement Protocol (STAMP), the proposed solution enables comprehensive NRP support in SR networks.

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 September 03, 2025.

Table of Contents

1. Introduction

[RFC9543] provides the definition of IETF network slice for use within the IETF and discusses the general framework for requesting and operating IETF Network Slices, their characteristics, and the necessary system components and interfaces. It also introduces the concept Network Resource Partition (NRP), which is a subset of the resources and associated policies in the underlay network.

Using OAM tools enables real-time monitoring of the operational status of network slices, allowing for quick detection and localization of faults. When a node or link within a network slice experiences a failure, OAM tools can promptly issue alerts, assisting network administrators in taking swift corrective action to minimize service downtime. Therefore, the use of OAM tools in an NRP network is crucial for ensuring the availability and performance of network slice resources. This not only enhances user experience but also improves the overall efficiency and stability of the network.

[RFC8402] describes Segment Routing Architecture and its instantiation in two data planes: MPLS and IPv6.

For different data plane types, existing OAM tools can be utilized for inspection. Existing OAM tools typically include Ping, Traceroute, BFD, and STAMP. [RFC9259] explains how to perform OAM when the underlying network uses SRv6. [RFC8029] describes how to Detect MPLS Data-Plane Failures in MPLS networks.

This document continues to employ these existing OAM mechanisms to monitor NRP resources. Specifically, it describes the OAM mechanisms for inspecting NRP resources in scenarios where the data plane is based on MPLS networks and SRv6 networks. However, the OAM methods for inspecting NRP resources outlined in this document are not limited to these specific data plane types.

This document outlines how to utilize existing OAM tools to monitor the operational status of NRP resources and quickly detect and pinpoint faults within NRP resources.

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. OAM Mechanisms

During the process of using existing OAM mechanisms to check the operational status of NRP resources, the OAM initiator needs to carry the NRP-ID in the data plane of the inspection packets.

Intermediate equipment and OAM End Points need to check the availability of NRP resources when receiving OAM packets with an NRP-ID. If the NRP resources are unavailable, they should respond to the OAM initiator with an error message, indicating that the NRP resources are unavailable.

This document adopts these existing methods of carrying the NRP-ID in the data plane to perform OAM operations within NRP networks. The specific mechanisms for carrying the NRP-ID in the data plane are outside the scope of this document. Based on different underlying networks, this document describes how to use OAM tools to monitor NRP resources by carrying the NRP-ID during OAM operations.

Building on the aforementioned aspects, using existing OAM mechanisms for underlay network operations and existing mechanisms for carrying the NRP-ID in the data plane, this document will describe how to use OAM tools to monitor the operational status of NRP resources within NRP networks.

2.1. PING

When performing a PING operation, the initiator sends a PING request. To support the probing of NRP resources, NRP information is carried in the data layer. Intermediate nodes inspect the NRP resources. If the request packet can be forwarded to the control plane, the response packet can include an error code to notify the initiator of an "NRP resource unavailable" error. However, if the packet cannot be forwarded to the control plane, the request packet is simply dropped, and the initiator cannot obtain specific error information.

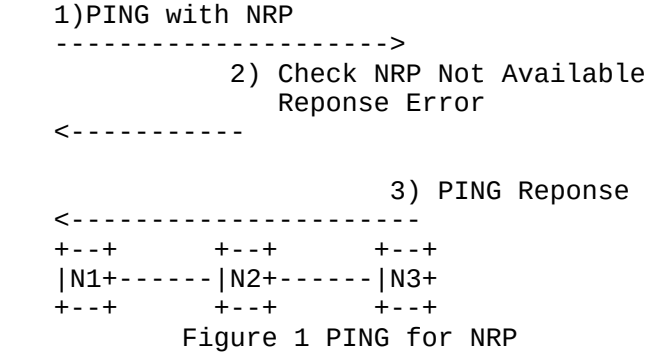


Figure 1 PING for NRP

Process of PING for NRP:

1) The initiator of the PING request includes the NRP-ID in the data layer when sending the PING request.

   [RFC8029]When the data layer is MPLS, the PING Request is an MPLS Echo Request message.

   [RFC9259]When the data layer is SRv6, the PING Request is an ICMPv6 message encapsulated with an SRH header.

2) The intermediate node or End Point first checks if the NRP resources are available when processing a Ping Request. If they are not available, it responds with a Ping Response, indicating the Error as "NRP resources unavailable".

   For MPLS networks, it is necessary to extend the Return Codes carried in the MPLS Echo Reply(IANA 6.1).

   For SRv6 networks, the Error Code in the ICMPv6 Reply needs to be extended(IANA 6.2).

3) If the check passes, the End Point will respond with a normal PING Response.

2.2. TRACEROUTE

When performing a TRACEROUTE operation, the TRACEROUTE initiator sends request packets toward the destination node by incrementally increasing the TTL value. To support the probing of NRP resources, NRP information is carried in the data layer. Each intermediate node first checks the availability of NRP resources before inspecting the TTL. If the resources are unavailable, the node responds with an error message indicating resource unavailability. In both MPLS and SRv6 networks, the packets used for TRACEROUTE are the same as those used for PING. When NRP resources are unavailable, the error codes used are also identical to those used in PING operations

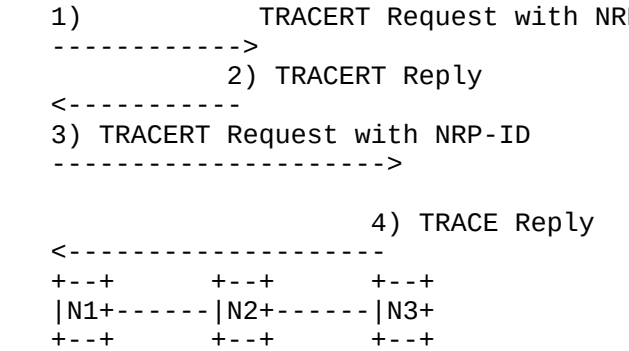


Figure 2 Traceroute for NRP

Process of Traceroute for NRP:

1) The initiator of the Traceroute request includes the NRP-ID in the data layer when sending the Traceroute request.

   [RFC8029]When the data layer is MPLS, the Traceroute Request is an MPLS Echo Request message with TTL 1 to n increase.

   [RFC9259]When the data layer is SRv6, the Traceroute Request is an ICMPv6 message encapsulated with an SRH header with TTL 1 to n increase.

2) The intermediate node or End Point first checks if the NRP resources are available when processing a Traceroute Request. If they are not available, it responds with a Traceroute Response, indicating the Error as "NRP resources unavailable". The error code for expansion should be the same as PING.

3) If the check passes, the process proceeds with a normal Traceroute, performing hop-by-hop detection of the path to the End Point until the Traceroute process is completed, and the detection results are outputted.

2.3. BFD

[RFC5880][RFC5881] provides a detailed description of the BFD protocol.

By establishing a BFD session between the head node and the tail node, BFD packets are exchanged regularly to quickly detect the reachability of the route between these nodes.

For NRP support, the BFD packets are encapsulated with an NRP-ID to probe the reachability of the corresponding route within this NRP.

When intermediate nodes and the tail node receive a BFD packet, they must check the availability of NRP resources. If the resources are unavailable, the received BFD packet should be discarded.

It is important to note that the identifier for a BFD session typically does not include the NRP-ID. Instead, it consists of [LD (Local Discriminator), RD (Remote Discriminator), source address, destination address]. When establishing BFD sessions with identifiers across different NRPs, it is impossible to distinguish between these sessions based solely on the identifiers. Therefore, in such cases, distinct LD and RD values must be configured for each NRP, enabling differentiation between BFD sessions based on [LD (Local Discriminator), RD (Remote Discriminator), source address, destination address].

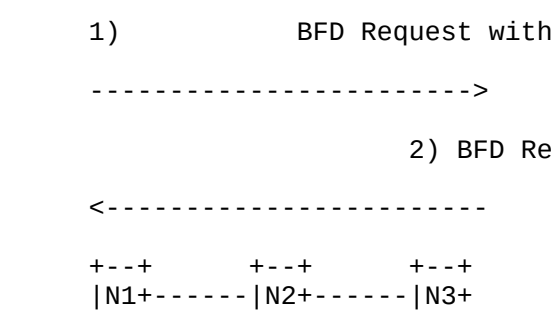


Figure 3 BFD for NRP

Process of BFD for NRP:

1) The session initiator includes the NRP-ID in the data layer.

2) The responder first checks the availability of NRP resources. If the NRP resources are unavailable, it does not respond. Otherwise, it replies with a BFD detection response.

2.4. STAMP

[RFC8762] describes the implementation process of the STAMP protocol.

When sending a STAMP Request, the NRP information to be inspected is carried in the data layer. Intermediate nodes processing the STAMP Request perform an NRP resource check. If the request packet can be forwarded to the control plane, the STAMP response packet can include an error code to notify the initiator of an "NRP resource unavailable" error. However, if the packet cannot be forwarded to the control plane, the request packet is simply dropped, and the initiator cannot obtain specific error information.

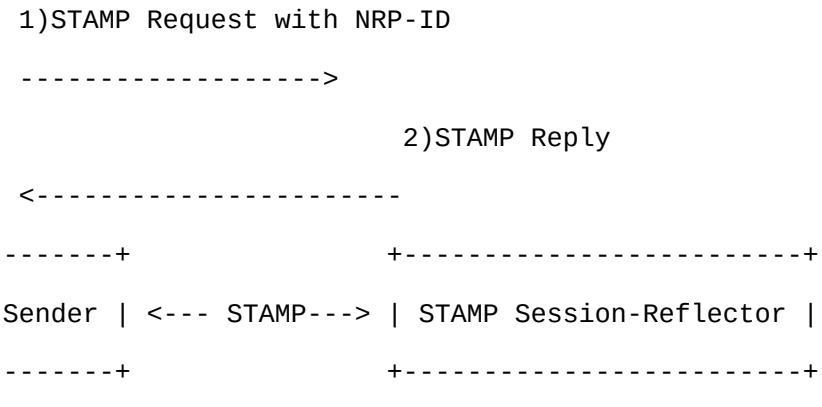


Figure 4 STAmp for NRP

1) The STAMP Session-Sender transmits STAMP packets, carrying the NRP-ID in the data layer.

2) The STAMP Session-Reflector first checks the availability of NRP resources. If the resources are unavailable, it responds with an "NRP resources unavailable" error. Otherwise, it proceeds to reply with a normal STAMP response.

3. Round-trip path consistency

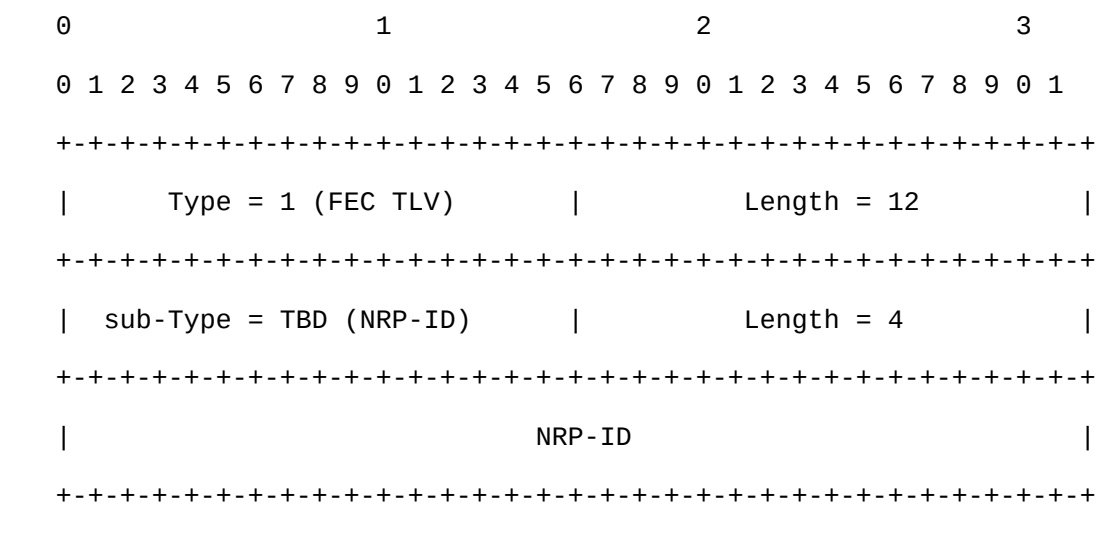
For scenarios where OAM checks are needed for the return path of NRP resources, there are two methods:

The first method involves the head node carrying the return path’s NRP-ID in the control plane. This return NRP-ID is sent to the end-point via the OAM message. The end-point retrieves the return NRP-ID

from the control plane data and uses it as the NRP-ID for the data plane of the response message.

[RFC8029]When performing OAM operations in an MPLS network, the return path can be specified by including the Reverse-path Target FEC Stack TLV in Request message. Based on this, this document extends the Reverse-path Target FEC Stack TLV by adding a return NRP-ID field.

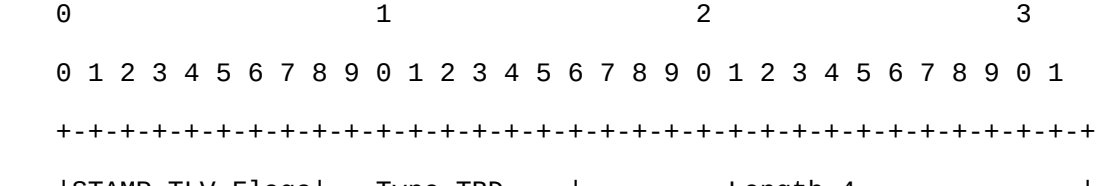
[RFC6426]The format of the Reverse-path Target FEC Stack TLV is the same as that of the Target FEC Stack TLV defined in [RFC4379].A Target FEC Stack is a list of sub-TLVs, The new sub-TLV "NRP-ID" is defined in this document, the format is as follow:



Sub-Type Refer to IANA 4.3.

[RFC9503]When performing STAMP, the Return Path can be included in the request message to specify the return path. Based on this, this document extends the STAMP Return Path Sub-TLV to carry the NRP-ID within the Return Path.

The format of Return NRP-ID sub-TLV in Return Path Sub-TLVs:



Sub-Type Refer to IANA 4.4.

The second method involves pre-configuring the return NRP-ID at the end-point by an administrator. When the end-point receives an OAM message from the head point, it uses the pre-configured return NRP-ID as the NRP-ID for the data plane of the response message.The specific implementation of this method is not within the scope of this document.

4. UseCase

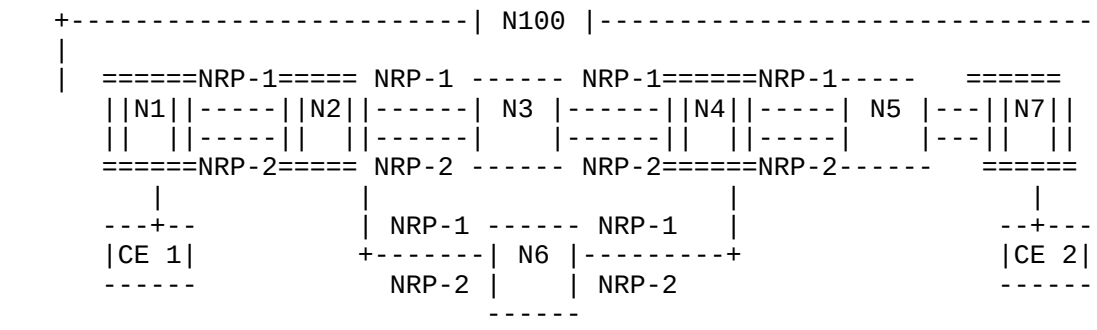


Figure 5 NRP network diagram

As illustrated In the reference topology of Figure 1,

Node j has a IPv6 loopback address 2001:db8:L:j::/128.

A SID at node j with locator block 2001:db8:K::/48 and function U is represented by 2001:db8:K:j:U::.

Node j has a IPv4 loopback address 192.168.j.1/32

A LABEL at node j is 1j000.

Node N100 is a controller.

The IPv6 address of the Link between node i and j at the NRP-ID is represented as 2001:db8:i:j:nrp::

4.1. PING in MPLS network

An example of MPLS Ping success:

> ping 15000 via label-stack 12000, 14000, NRP-ID: 1, Ret NRP-ID: 2

Sending 5, 100-byte MPLS Echos to 192.168.5.2, timeout is 2 seconds:

!!!!!

Success rate is 100 percent (5/5), round-trip min/avg/max = 0.625 /0.749/0.931 ms

An example of MPLS Ping failure due to NRP resource unavailability:

> ping 15000 via label-stack 12000, 14000, NRP-ID: 1, Ret NRP-ID: 2

Reply to request 2 (1 ms). Return Code: ’N’

Reply to request 3 (1 ms). Return Code: ’N’

Reply to request 4 (1 ms). Return Code: ’N’

Reply to request 3 (1 ms). Return Code: ’N’

Reply to request 4 (1 ms). Return Code: ’N’

Success rate is 0 percent (0/5), round-trip min/avg/max = 1/1/1 ms

Error code ’N’ indicates that the cause of the error is the unavailability of NRP resources. This explanation applies to the following examples as well and will not be reiterated.

4.2. TRACEROUTE in MPLS network

An example of MPLS traceroute success:

> traceroute 15000 via label-stack 12000, 14000, NRP-ID: 1, Ret-NRP-ID: 2

Tracing the route to 15000

1 192.168.2.1 [MPLS: Label 12000] 1.123 ms 1.045 ms 1.067 ms

2 192.168.4.1 [MPLS: Label 14000] 1.123 ms 1.045 ms 1.067 ms

2 192.168.5.1 [MPLS: Label 15000] 1.123 ms 1.045 ms 1.067 ms

An example of MPLS traceroute failure due to NRP resource unavailability:

> traceroute 15000 via label-stack 12000, 14000, NRP-ID: 1, Ret-NRP-ID: 2

Tracing the route to 15000

1 192.168.2.1 [MPLS: Label 12000] Return Code: ’N’

4.3. PING in SRv6 network

An example of SRv6 Ping success:

> ping 2001:db8:L:5:: via segment-list 2001:db8:K:2:X31::, 2001:db8:K:4:X52::, NRP-ID: 1, Ret NRP-ID: 2

!!!!!

Success rate is 100 percent (5/5), round-trip min/avg/max = 0.625 /0.749/0.931 ms

An example of SRv6 Ping failure due to NRP resource unavailability:

> ping 2001:db8:L:5:: via segment-list 2001:db8:K:2:X31::, 2001:db8:K:4:X52::, NRP-ID: 1, Ret NRP-ID: 2

Reply to request 2 (1 ms). Return Code: ’N’

Reply to request 3 (1 ms). Return Code: ’N’

Reply to request 4 (1 ms). Return Code: ’N’

Reply to request 3 (1 ms). Return Code: ’N’

Reply to request 4 (1 ms). Return Code: ’N’

4.4. TRACEROUTE in SRv6 network

An example of SRv6 traceroute success:

> traceroute 2001:db8:L:5:: via segment-list 2001:db8:K:2:X31::,
2001:db8:K:4:X52::, NRP-ID: 1, Ret-NRP-ID: 2

Tracing the route to 2001:db8:L:5::

1 2001:db8:2:1:21:: 0.512 msec 0.425 msec 0.374 msec DA:
2001:db8:K:2:X31::, SRH:(2001:db8:L:5::2, 2001:db8:L:5::1,
2001:db8:K:4:X52::1, 2001:db8:K:2:X31::1, SL=3)

2 2001:db8:3:2:31:: 0.721 msec 0.810 msec 0.795 msec DA:
2001:db8:K:4:X52::, SRH:(2001:db8:L:5::2, 2001:db8:L:5::1,
2001:db8:K:4:X52::1, 2001:db8:K:2:X31::1, SL=2)

3 2001:db8:4:3::41:: 0.921 msec 0.816 msec 0.759 msec DA:
2001:db8:K:4:X52::, SRH:(2001:db8:L:5::2, 2001:db8:L:5::1,
2001:db8:K:4:X52::1, 2001:db8:K:2:X31::, SL=1)

4 2001:db8:5:4::52:: 0.879 msec 0.916 msec 1.024 msec DA:
2001:db8:L:5::

An example of SRv6 traceroute failure due to NRP resource unavailability:

> traceroute 2001:db8:L:5:: via segment-list 2001:db8:K:2:X31::,
2001:db8:K:4:X52::, NRP-ID: 1, Ret-NRP-ID: 2

Tracing the route to 2001:db8:L:5::

1 2001:db8:2:1:21::, Return Code: ’N’

5. Security Considerations

This document does not impose any additional security challenges to be considered beyond the security threats described in [RFC4884], [RFC4443], [RFC0792], [RFC8754], and [RFC8986].

6. IANA Considerations

6.1. MPLS Reply Error Code

IANA is requested to allocated new Return Codes "Return Subcode" registry.

Value	Meaning
TBD	NRP resource unavailable

6.2. ICMPv6 Error Code

Type 3 - Destination Unreachable

Code: TBD. Description: "NRP resource unavailable".

6.3. NRP-ID sub TLV in Reverse-path Target FEC Stack TLV

Sub-Type	Length	Value Field
TBD	4	NRP-ID

6.4. STAMP Return Path Sub-TLV

IANA is requested to allocated new type Sub-TLV Types in the "Return Path Sub-TLV Types" registry.

Value	Description	Reference
TBD	Return Path NRP-ID	This Doc

7. References

7.1. Normative References

[RFC9259] Z. Ali, C. Filsfils, Cisco Systems, S. Matsushima, Softbank, D. Voyer, Bell Canada, M. Chen, Huawei,"Operations, Administration, and Maintenance (OAM) in Segment Routing over IPv6 (SRv6)", RFC 9259, DOI 10.17487/RFC9259, June 2022, <https://www.rfc-editor.org/info/rfc9259>.

[RFC9543] Farrel, A., Ed., Drake, J., Ed., Rokui, R., Homma, S., Makhijani, K., Contreras, L., and J. Tantsura, "A Framework for Network Slices in Networks Built from IETF Technologies", RFC 9543, DOI 10.17487/RFC9543, March 2024, <https://www.rfc-editor.org/info/rfc9543>.

[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>.

[RFC9503] R. Gandhi, C. Filsfils, Cisco Systems, M. Chen, Huawei, B. Janssens, Colt, R. Foote, Nokia, "Simple Two-Way Active Measurement Protocol (STAMP) Extensions for Segment Routing Networks", RFC 9503, DOI 10.17487/RFC9503, October 2023, <https://www.rfc-editor.org/info/rfc9503>.

[RFC8762] Mirsky, G., Jun, G., Nydell, H., and R. Foote, "Simple Two-Way Active Measurement Protocol", RFC 8762, DOI 10.17487/RFC8762, March 2020, <https://www.rfc-editor.org/info/rfc8762>.

[RFC8029] K. Kompella, Juniper Networks, Inc., G. Swallow, C. Pignataro, Ed., N. Kumar, Cisco, S. Aldrin, Google, M. Chen, Huawei, "Detecting Multiprotocol Label Switched (MPLS) Data-Plane Failures", RFC 8029, DOI 10.17487/RFC8029, March 2017, <https://www.rfc-editor.org/info/rfc8029>.

[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>.

[RFC8972] Mirsky, G., Xiao, M., Nydell, H., Foote, R., Masputra, A., and E. Ruffini, "Simple Two-way Active Measurement Protocol Optional Extensions", Work in Progress, Internet-Draft, draft-ietf-ippm-stamp-option-tlv-03, 21 February 2020, <https://tools.ietf.org/html/draft-ietf-ippm-stamp-option-tlv-03>.

7.2. Informative References

TBD

Acknowledgements

TBD

Authors' Addresses

Liyan Gong
China Mobile
China
Email: gongliyan@chinamobile.com

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