

Workgroup: spring
Internet-Draft:
draft-zzhang-spring-microtap-segment-01
Published: 13 March 2023
Intended Status: Standards Track
Expires: 14 September 2023
Authors: Z. Zhang R. Hoffman G. Bajwa
 Juniper Networks TELUS TELUS
 D. Voyer S. Zadok A. Wang L. Jalil
 Bell Canada Broadcom China Telecom Verizon
 S. Li
 Arrcus

MicroTap Segment in Segment Routing

Abstract

This document specifies a microTap segment that can be used to instruct a transit node to make a copy of a segment-routed packet and deliver it to a specified node for the purpose of network monitoring, trouble shooting, or lawful intercept.

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 14 September 2023.

Copyright Notice

Copyright (c) 2023 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. Specification](#)
 - [2.1. Signaling](#)
 - [2.1.1. OSPF Signaling](#)
 - [2.1.2. ISIS Signaling](#)
 - [2.1.3. BGP Signaling](#)
 - [2.2. Controller Signaling](#)
 - [2.2.1. BGP-LS](#)
 - [2.2.2. PCEP](#)
 - [2.3. Procedures](#)
 - [2.3.1. Optional IOM header](#)
- [3. Security Considerations](#)
- [4. IANA Assignments](#)
- [5. Contributors](#)
- [6. References](#)
 - [6.1. Normative References](#)
 - [6.2. Informative References](#)
- [Authors' Addresses](#)

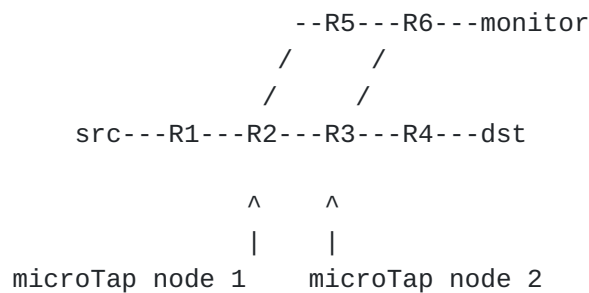
1. Introduction

Network operators may for various reasons benefit from the ability to tap packets at strategic locations within their respective networks. Segment routing [[RFC8402](#)] technology offers the ability to both simplify and improve the operational experience of deploying targeted packet tapping.

The tapping can be only for some random packets for monitoring purposes, so we use the term microTap and tap interchangeably in this document.

The introduction and strategic placement within a SID-list of one or more microTap SIDs can signal the desire to tap traffic at targeted points within the network without the need for explicit configuration on those nodes.

Consider an SR network in the following example diagram where traffic is steered along some paths by using a SID-list in the packets. For network debugging/monitoring purposes, the operator may at any time want for a certain node (e.g., R2 or R3) in the network to tap a copy of a packet to a monitor (e.g. connected to R6), while continue to forward the original packet along its path to the destination.



To make it very flexible and precise on specifying which packets to tap on what node and avoid the need to configure filters on the microTap node, a microTap SID can be inserted to the SID-list after a Node-SID (for the microTap node) or an Adjacency-SID (that leads to the microTap node). When the microTap SID becomes the current active SID, the node does the following:

- *Replicate the packet, and send the copy to the remote monitor

- *Pop the microTap SID off the original packet and continue forwarding

There could be multiple monitors. A microTap SID is associated with a particular monitor (vs. a microTap node). In the above example, there could be another monitor attached to R5. In that case, there would be two microTap SIDs - one for the monitor attached at R5 (say microTap SID S5) and one for the monitor attached at R6 (say microTap SID S6). The monitor could be a separate server attached to an interface on R5 or R6, or could be an internal service entity on R5 or R6 (which can be viewed as connected via an internal interface).

If S5 becomes the active SID in a packet arriving at R2, R2 will tap the packet to R5, by imposing R5's node SID label on top of S5. When the tapped copy arrives at R5, R5 knows that the packet should be sent to the internal or external monitor (because S5, which R5 advertises, becomes the active SID). Similarly, if S6 becomes the active SID in a packet arriving at R3, R3 will tap the packet to R6, by imposing R6's node SID label on top of S6. In case of SRv6, a separate IPv6 header is used to send the packet to the router to which the monitor is attached.

A microTap SID is advertised by the router that hosts the monitor. It should only become the active SID in a packet arriving at the desired microTap node or the advertising/owning node. A node supporting microTap functionality advertises its ability to do so, so that incapable nodes will never see a microTap SID as the active SID in a packet.

The SID-list may contain multiple microTap SIDs that may or may not be adjacent in the list. For nonadjacent microTap SIDs, different

nodes will tap to the same or different monitors (depending on the value of microTap SIDs). For adjacent microTap SIDs in the list, they are likely for different monitors - for the "continue forwarding" part of the first microTap SID, the second microTap SID becomes active segment, leading to the second microTap operation.

2. Specification

2.1. Signaling

A node (e.g. R2/R3) supporting microTap function advertise its capability to other nodes.

A node (e.g. R5/R6) hosting a monitor is provisioned with a microTap SID allocated from the SRGB. The microTap SID is advertised to other nodes.

A microTap SID MUST be associated with only one specific monitor.

If the same microTap SID value is advertised by more than one node, it MUST be treated by a receiving node as an error and ignored, and MUST NOT be used in the SID-List of a packet.

SRv6 related signaling details will be added in future revisions.

2.1.1. OSPF Signaling

This document defines a new TLV for the advertisement of a microTap SID (from a node hosting a monitor) and an existing TLV is leveraged for the advertisement of tapping capability (from a microTap node).

2.1.1.1. MicroTap-SID TLV

The microTap SID is advertised in a newly defined MicroTap-SID Sub-TLV that mimics the Prefix SID Sub-TLV as defined in Section 5 of [\[RFC8665\]](#):

2.1.2. ISIS Signaling

ISIS signaling is similar to OSPF, as specified in the following sections.

2.1.2.1. MicroTap-SID

The microTap SID is advertised in a newly defined MicroTap-SID Sub-TLV that mimics the Prefix SID Sub-TLV as defined in Section 2.1 of [\[RFC8667\]](#):

[illegible]

where:

Type: To be assigned by IANA.

Length: 5 or 6 depending on the size of the SID (described below)

Flags: 1-octet field. Currently no flags are defined.

Algorithm: the router may use various algorithms when calculating reachability to other nodes or to prefixes attached to these nodes. Algorithm identifiers are defined in Section 3.2. Examples of these algorithms are metric-based Shortest Path First (SPF), various sorts of Constrained SPF, etc. The Algorithm field of the Prefix-SID contains the identifier of the algorithm the router uses to compute the reachability of the prefix to which the Prefix-SID is associated.

At origination, the Prefix-SID Algorithm field MUST be set to 0 or any value advertised in the SR-Algorithm sub-TLV.

A router receiving a Prefix-SID from a remote node and with an algorithm value that such remote node has not advertised in the SR-Algorithm sub-TLV MUST ignore the Prefix-SID sub-TLV.

SID/Index/Label: : Currently a 4-octet index defining the offset in the Segment Routing Global Block (SRGB) advertised by his router. In the future the flags field may change the definition of this definition of this field.

The MicroTap-SID Sub-TLV MAY appear where a Prefix-SID Sub-TLV is included to advertises a node SID.

2.1.2.2. Tapping Capability

Similar to OSPF, a new flag T in the Flags field of the Prefix/Adjacency-SID Sub-TLV indicates that a MicroTap SID is allowed to follow the prefix/adjacency SID in a packet:

```

      0  1  2  3  4  5  6  7
+---+---+---+---+---+---+---+
|   |   |   |   |   |   | T |
+---+---+---+---+---+---+---+

```

2.1.3. BGP Signaling

2.1.3.1. MicroTap-SID

A new MicroTap-SID TLV is defined to advertise a microTap SID. It has the same encoding as the Label-Index TLV except with a different type. The following is copied verbatim from [\[RFC8669\]](#):

```

      0                               1                               2                               3
      0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|           Type           |           Length           | RESERVED |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|           Flags           |           Label Index           |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|           Label Index           |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

where:

Type: To be assigned by IANA.

Length: 7, the total length in octets of the value portion of the TLV.

RESERVED: 8-bit field. It MUST be clear on transmission and MUST be ignored on reception.

Flags: 16 bits of flags. None are defined by this document. The Flags field MUST be clear on transmission and MUST be ignored on reception.

Label Index: 32-bit value representing the index value in the SRGB space.

A MicroTap-SID TLV MAY be included in the BGP Prefix-SID attribute.

2.1.3.2. Tapping Capability

A 'T' flag is defined for the existing Originator SRGB TLV's Flags field to indicate that the originator supports microTapping functionality. Exact bit position for the flag is to be assigned by IANA and registered in the "BGP Prefix-SID Originator SRGB TLV Flags" registry.

2.2. Controller Signaling

A controller needs to know about the nodes (e.g. R2/R3) that support tapping function, and the nodes (e.g. R5/R6) hosting a monitor & relevant microTap SID. This information is advertised to the controller by the link-state routing protocols (ISIS and OSPF) or BGP-LS. The signaling for OSPF and ISIS has been covered in the previous sections of this document. This section covers signaling for BGP-LS and PCEP.

2.2.1. BGP-LS

This document defines a new TLV for the advertisement of a microTap SID (from a node hosting a monitor) and an existing TLV is leveraged for the advertisement of tapping capability (from a microTap node).

2.2.1.1. MicroTap SID

The microTap SID is advertised in a newly defined MicroTap-SID TLV that mimics the Prefix SID TLV as defined in Section 2.3.1 of [\[RFC9085\]](#):


```

0          1          2          3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                                     |                                     |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|      Flags      |  Algorithm  |      Reserved      |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                                     SID/Index/Label (variable)                                     //
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

Where:

Type: To be assigned by IANA

Length: Variable. 7 or 8 octets depending on the label or index encoding of the SID.

Flags: 1-octet value that should be set as:

- * IS-IS MicroTap-SID flags as defined in Section 2.1.2.1.
- * OSPFv2 MicroTap-SID flags as defined in Section 2.1.1.1.
- * OSPFv3 MicroTap-SID flags as defined in Section 2.1.1.1.

Algorithm: 1-octet value identifies the algorithm. The semantics of the algorithm are described in Section 3.1.1 of {{RFC8402}}.

Reserved: 2 octets that MUST be set to 0 and ignored on receipt.

SID/Index/Label:

IS-IS: Label or index value as defined in Section 2.1.2.1.

OSPFv2: Label or index value as defined in Section 2.1.1.1.

OSPFv3: Label or index value as defined in Section 2.1.1.1.

The Flags and, as an extension, the SID/Index/Label fields of this TLV are interpreted according to the respective underlying IS-IS, OSPFv2, or OSPFv3 protocol. The Protocol-ID of the BGP-LS Prefix NLRI is used to determine the underlying protocol specification for parsing these fields.

The MicroTap-SID TLV MAY appear where a Prefix-SID TLV advertises a node SID.

2.2.1.2. Tapping Capability

The Flags of Prefix/Adjacency-SID TLV are interpreted according to the respective underlying IGP specification. The new flag T in the Flags field of the Prefix/Adjacency-SID TLV indicates that a MicroTap SID is allowed to follow the prefix/adjacency SID in a packet.

2.2.2. PCEP

An SR-TE path consists of one or more SIDs and may contain one or more microTap SIDs. The SR-TE path information is exchanged between the PCE and PCC in ERO and RRO subobjects. The SR-ERO subobject and SR-RRO subobject defined in [\[RFC8664\]](#) are used to carry a SID which can be a microTap SID.

2.3. Procedures

The node hosting a monitor treats a microTap SID that it advertises as an adjacency SID. In other words, it sets up its forwarding state for the microTap SID such that packets with the microTap SID as current active SID will be sent to the monitor (after popping the microTap SID). It is the responsibility of the monitor to parse the packet (including the remaining SID-list).

A node supporting microTap functionality sets up its forwarding state for each microTap SID that it receives, such that packets with the microTap SID as current active SID are processed as following:

- *Make a copy and send it to the advertising node of the microTap SID. In case of SR-MPLS, this is done by imposing the advertising node's node SID (optionally after imposing the node SID of the microTap node so that the monitor knows the microTap node). In case of SRv6, this is done by imposing an outer IPv6 encapsulation with the destination address being the advertising node's address.

- *Forward the original packet after popping the microTap SID

If a node does not support microtapping but does recognize the microtap SID signaling, the forwarding behavior for the SID is simply pop on that node. This is to safeguard the situation in case the node received a packet with the active SID being a microtap SID.

The ingress node may add microTap SIDs to the SID-list of a packet based on its monitoring/debugging needs or based on SR policies programmed from a controller.

A microTap SID MUST not be placed in the SID-list after a node or adjacency SID that is for or leads to a node that does not advertise

microTap capability. Otherwise, the packet with that SID-list will be discarded by the node.

In case of SRv6, the microTap SID and its preceding node SID MAY be merged into a single IPv6 address in SRH: the locator part identifies the microTap SID and the function part is the 3-octet or 4-octet microTap SID.

2.3.1. Optional IOAM header

As replicated packets traverse the network from the microtap node to the monitor nodes, packet loss, packet reordering and buffering can occur. To allow packet analysis equipment that receives these replicated packets to accurately analyze the replicated packet flow, additional information is needed in the replicated packet header to recreate the original conditions of the flow.

RFC9197] defines a header with data fields well suited for this purpose. IOAM includes timestamp data, indicating the arrival time the replicated packet was received at the microtap node. This timestamp can be used to reproduce accurate inter-packet gaps during packet analysis. IOAM also includes a sequence number, indicating the order of replicated packets received by the microtap node. This sequence number can be used by the packet analysis equipment to reorder packets, remove duplicated packets, and to alarm on the condition that replicated packets were lost in transit.

The microTap node MAY include an IOAM header in the replicated packet with following fields:

- *Timestamp Seconds

- *Timestamp Fraction

- *64-bit sequence number

It is RECOMMENDED that all nodes that perform microtap packet replication be Time of Day (ToD) synchronized via Precision Time Protocol (PTP) for the most accurate recreation of packet conditions during analysis.

The added IOAM header is Edge-to-Edge Option-Type, and in addition to possible IOAM header already present when the packet arrives at the microtap node. In case of MPLS, the added IOAM header is an MPLS extension header [[I-D.song-mpls-extension-header](#)] that follows the Node SID of the node that originated the microtap SID. The extension header is followed by the original label stack and its OUL field (Original Upper Layer protocol type) MUST be set to MPLS. In other words, there may be two label stacks in the packet arriving at the node hosting the monitoring station.

If MTU is a concern, the original label stack (except the microTap SID) and extension headers MAY be removed.

3. Security Considerations

To be added.

4. IANA Assignments

To be added.

5. Contributors

The following also contributed to this document:

Peter Van Oene
Juniper Networks
pvanoene@juniper.net

6. References

6.1. Normative References

- [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>>.
- [RFC8665] Psenak, P., Ed., Previdi, S., Ed., Filsfils, C., Gredler, H., Shakir, R., Henderickx, W., and J. Tantsura, "OSPF Extensions for Segment Routing", RFC 8665, DOI 10.17487/RFC8665, December 2019, <<https://www.rfc-editor.org/info/rfc8665>>.
- [RFC8667] Previdi, S., Ed., Ginsberg, L., Ed., Filsfils, C., Bashandy, A., Gredler, H., and B. Decraene, "IS-IS Extensions for Segment Routing", RFC 8667, DOI 10.17487/RFC8667, December 2019, <<https://www.rfc-editor.org/info/rfc8667>>.
- [RFC8669] Previdi, S., Filsfils, C., Lindem, A., Ed., Sreekantiah, A., and H. Gredler, "Segment Routing Prefix Segment Identifier Extensions for BGP", RFC 8669, DOI 10.17487/RFC8669, December 2019, <<https://www.rfc-editor.org/info/rfc8669>>.
- [RFC9085] Previdi, S., Talaulikar, K., Ed., Filsfils, C., Gredler, H., and M. Chen, "Border Gateway Protocol - Link State (BGP-LS) Extensions for Segment Routing", RFC 9085, DOI

10.17487/RFC9085, August 2021, <<https://www.rfc-editor.org/info/rfc9085>>.

6.2. Informative References

[RFC8664] Sivabalan, S., Filsfils, C., Tantsura, J., Henderickx, W., and J. Hardwick, "Path Computation Element Communication Protocol (PCEP) Extensions for Segment Routing", RFC 8664, DOI 10.17487/RFC8664, December 2019, <<https://www.rfc-editor.org/info/rfc8664>>.

[I-D.song-mpls-extension-header] Song, H., Zhou, T., Andersson, L., Zhang, Z. J., and R. Gandhi, "MPLS Network Actions using Post-Stack Extension Headers", Work in Progress, Internet-Draft, draft-song-mpls-extension-header-11, 15 October 2022, <<https://datatracker.ietf.org/doc/html/draft-song-mpls-extension-header-11>>.

Authors' Addresses

Zhaohui Zhang
Juniper Networks

Email: zzhang@juniper.net

Ryan Hoffman
TELUS

Email: ryan.hoffman@telus.com

Gurminderjit Bajwa
TELUS

Email: gurminderjit.bajwa@telus.com

Daniel Voyer
Bell Canada

Email: daniel.voyer@bell.ca

Shay Zadok
Broadcom

Email: shay.zadok@broadcom.com

Aijun Wang
China Telecom

Email: wangaj3@chinatelecom.cn

Luay Jalil
Verizon

Email: luay.jalil@verizon.com

Shengtao Li
Arcus

Email: shen@arcus.com