Path Segment in MPLS Based Segment Routing Network

Abstract

A Segment Routing (SR) path is identified by an SR segment list. A sub-set of segments from the segment list cannot distinguish one SR path from another as they may be partially congruent. SR path identification is a pre-requisite for various use-cases such as Performance Measurement (PM), and end-to-end 1+1 path protection.

In SR for MPLS data plane (SR-MPLS), an Egress node cannot determine on which SR path a packet traversed the network from the label stack because the segment identifiers are removed from the label stack as the packet transits the network.

This document defines Path Segment to identify an SR path on the egress node of the path.

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 13 April 2024.
1. Introduction

Segment Routing (SR) [RFC8402] leverages the source-routing paradigm to steer packets from a source node through a controlled set of instructions, called segments, by prepending the packet with an SR
header. In the MPLS data plane SR-MPLS [RFC8660] the SR header is instantiated through a label stack.

In an SR-MPLS network, when a packet is transmitted along an SR path, the labels in the MPLS label stack will be swapped or popped. The result of this is that no label or only the last label may be left in the MPLS label stack when the packet reaches the egress node. Thus, the egress node cannot use the SR label stack to determine along which SR path the packet came.

However, to support various use-cases in SR-MPLS networks, such as Performance Measurement (PM) Section 3.1, bidirectional path Section 3.2, and end-to-end 1+1 path protection (Live-Live case) Section 3.3, the ability to implement path identification on the egress node is a prerequisite.

Therefore, this document defines a new segment type, referred to herein as a Path Segment. A Path Segment is defined to uniquely identify an SR path on the egress node of the path. It MAY be used by the egress node for path identification. Note that, per-path state will be maintained in the egress node due to the requirements in the aforementioned use cases, though in normal cases that the per-path state will be maintained in the ingress node only.

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.

1.2. Abbreviations and Terms

DM: Delay Measurement.

LM: Loss Measurement.

MPLS: Multiprotocol Label Switching.

MSD: Maximum SID Depth.

PM: Performance Measurement.

PSID: Path Segment ID.

SID: Segment ID.

SL: Segment List.
SR: Segment Routing.

SRLB: SR Local Block

SR-MPLS: Instantiation of SR on the MPLS data plane.

SR path: A SR path is a path described by a Segment-List.

Sub-Path: A sub-path is a part of the a path, which contains a subset of the nodes and links of the path.

2. Path Segment

A Path Segment is a Local Segment which uniquely identifies an SR path on the egress node. A Path Segment Identifier (PSID) is a single label that is assigned from the Segment Routing Local Block (SRLB) [RFC8402] of the egress node of an SR path.

A PSID is used to identify a Segment List. However, one PSID can be used to identify multiple Segment Lists in some use cases if needed. For example, one single PSID MAY be used to identify some or all Segment lists in a Candidate path or an SR policy, if an operator would like to aggregate these Segment Lists in operation.

When a PSID is used, the PSID can be inserted at the ingress node and MUST immediately follow the last label of the SR path, in other words, inserted after the routing segment (adjacency/node/prefix segment) pointing to the egress node of the SR path. Therefore, a PSID will not be the top label in the label stack when received on an intermediate node of the associated path, but it can be the top label in the label stack on the penultimate node.

The value of the TTL field in the MPLS label stack entry containing a PSID can be set to any value except 0. If a PSID is the bottom label, the S bit MUST be set.

The egress node MUST pop the PSID. The egress node MAY use the PSID for further processing. For example, when performance measurement is enabled on the SR path, it can trigger packet counting or timestamping.

The addition of the PSID will require the egress to read and process the PSID label in addition to the regular processing (such as a below MPLS label or the MPLS payload). This additional processing may have an impact on forwarding performance.

Generic Associated Label (GAL) MAY be used for Operations, Administration and Maintenance (OAM) in MPLS networks. As per [RFC5586], when GAL is used, the ACH appears immediately after the bottom of the label stack.
The SR path computation needs to know the Maximum SID Depth (MSD) that can be imposed at the ingress node of a given SR path [RFC8664]. This ensures that the SID stack depth of a computed path does not exceed the number of SIDs the node is capable of imposing. As per [RFC8491] the MSD signals the total number of MPLS labels that can be imposed, where the total number of MPLS labels includes the PSID.

The label stack with Path Segment is shown in Figure 1:

```
+--------------------+
|                   ...
+--------------------+
|    Label 1        |
+--------------------+
|    Label 2        |
+--------------------+
|                   ...
+--------------------+
|    Label n        |
+--------------------+
|    PSID           |
+--------------------+
|        Payload     |
+--------------------+
```

Figure 1: Label Stack with Path Segment

Where:

*The Labels 1 to n are the segment label stack used to direct how to steer the packets along the SR path.

*The PSID identifies the SR path in the context of the egress node of the SR path.

Signaling of the PSID between the egress, ingress and possibly a centralized controller is out of the scope of this document.

2.1. Equal-Cost Multipath Considerations

If Entropy Label(EL) is also used on the egress node, as per [RFC6790] the Entropy label Indicator (ELI) and Entropy Label (EL) would be placed before the tunnel label and hence does not interfere with the PSID which is placed below.

It is worthy to note that in case of ECMP, with or without the use of EL, the SR packets may be forwarded over multiple paths. In this case, the SID list cannot directly reflect the actual forwarding
path and the PSID can only identify the SID list rather than the actual forwarding path.

Also, similar to Synonymous Flow Labels (SFL) [RFC8957], the introduction of an PSID to an existing flow may cause that flow to take a different path through the network under conditions of Equal-Cost Multipath (ECMP). This, in turn, may invalidate certain uses of the PSID, such as performance measurement applications. Therefore, the considerations as per section 5 in [RFC8957] of SFL also apply to PSID in implementation.

3. Use cases

This section describes use cases which can leverage the Path Segment.

3.1. Path Segment for Performance Measurement

As defined in [RFC7799], performance measurement can be classified into Passive, Active, and Hybrid measurement. Since Path Segment is encoded in the SR-MPLS Label Stack as shown in Figure 1, existing implementation on the egress node can be leveraged for measuring packet counts using the incoming SID (the PSID).

For Passive performance measurement, path identification at the measuring points is the pre-requisite. Path Segment can be used by the measuring points (e.g., the ingress and egress nodes of the SR path or a centralized controller) to correlate the packet counts and timestamps from the ingress and egress nodes for a specific SR path, then packet loss and delay can be calculated for the end-to-end path, respectively.

Path Segment can also be used for Active performance measurement for an SR path in SR-MPLS networks for collecting packet counters and timestamps from the egress node using probe messages.

Path Segment can also be used for In-situ OAM for SR-MPLS to identify the SR Path associated with the in-situ data fields in the data packets on the egress node.

Path Segment can also be used for In-band PM for SR-MPLS to identify the SR Path associated with the collected performance metrics.

3.2. Path Segment for Bidirectional SR Path

In some scenarios, for example, mobile backhaul transport networks, there are requirements to support bidirectional paths, and the path is normally treated as a single entity. Forward and reverse directions of the path have the same fate, for example, failure in one direction will result in switching traffic at both directions.
MPLS supports this by introducing the concepts of co-routed bidirectional LSP and associated bidirectional LSP [RFC5654].

In the current SR architecture, an SR path is a unidirectional path [RFC8402]. In order to support bidirectional SR paths, a straightforward way is to bind two unidirectional SR paths to a single bidirectional SR path. Path Segments can then be used to identify and correlate the traffic for the two unidirectional SR paths at both ends of the bidirectional path.

The mechanism of constructing bidirectional path using path segment is out of the scope of this document and has been described in several documents, such as [I-D.ietf-pce-sr-bidir-path] and [I-D.ietf-idr-sr-policy-path-segment].

3.3. Path Segment for End-to-end Path Protection

For end-to-end 1+1 path protection (i.e., Live-Live case), the egress node of the path needs to know the set of paths that constitute the primary and the secondaries, in order to select the primary path packets for onward transmission, and to discard the packets from the secondaries [RFC4426].

To do this in Segment Routing, each SR path needs a path identifier that is unique at the egress node. For SR-MPLS, this can be the Path Segment label allocated by the egress node.

There then needs to be a method of binding this SR path identifiers into equivalence groups such that the egress node can determine for example, the set of packets that represent a single primary path. This equivalence group can be instantiated in the network by an controller using the Path Segments of the SR paths.

3.4. Nesting of Path Segments

Binding SID (BSID) [RFC8402] can be used for SID list compression. With BSID, an end-to-end SR path in a trusted domain can be split into several sub-paths, each sub-path is identified by a BSID. Then an end-to-end SR path can be identified by a list of BSIDs, therefore, it can provide better scalability.

BSID and PSID can be combined to achieve both sub-path and end-to-end path monitoring. A reference model for such a combination in (Figure 2) shows an end-to-end path (A->D) in a trusted domain that spans three sub-domains (Access, Aggregation and Core domain) and consists of three sub-paths, one in each sub-domain (sub-path (A->B), sub-path (B->C) and sub-path (C->D)). Each sub-path is associated with a BSID and a s-PSID.
The SID list of the end-to-end path can be expressed as <BSID1, BSID2, ..., BSIDn, e-PSID>, where the e-PSID is the PSID of the end-to-end path. The SID list of a sub-path can be expressed as <SID1, SID2, ..., SIDn, s-PSID>, where the s-PSID is the PSID of the sub-path.

Figure 2 shows the details of the label stacks when PSID and BSID are used to support both sub-path and end-to-end path monitoring in a multi-domain scenario.

4. Security Considerations

A Path Segment in SR-MPLS is a local label similar to other labels/Segment, such as a VPN label, defined in MPLS and SR-MPLS. The data plane processing of a PSID is a local implementation of an ingress node, or an egress node, which follows the same logic of existing MPLS dataplane. As per definition of PSID, only the egress node and the ingress node of the associated path will learn the information of PSID. The intermediate nodes of this path will not learn it.

A PSID may be used on an ingress node that is not the ingress of the associated path, if the associated label stack with PSID is part of a deeper label stack which represents a longer path. For example the case described in Section 3.4 and the related BSID is not used while the original label stack of sub-path is inserted as a part of the
whole label stack. In this case, the PSID must be distributed in a trusted domain under the considerations defined in Section 8.1 of [RFC8402].

A Path Segment is used within an SR-MPLS trusted domain [RFC8402] and must not leak outside the domain, therefore no new security threats are introduced comparing to current SR-MPLS. As per [RFC8402], SR domain boundary routers MUST filter any external traffic destined to a label associated with a segment within the trusted domain, this applies to Path Segment as well. Other security considerations of SR-MPLS, described in Section 8.1 of [RFC8402] applies to this document.

The distribution of a PSID from an egress nodes to an ingress nodes is performed within an SR trusted domain, and it is out of the scope of this document. The details of the mechanism and related security considerations will be described in other documents.

5. Implementation Status

[Note to the RFC Editor - remove this section before publication, as well as remove the reference to [RFC7942].]

This section records the status of known implementations of the protocol defined by this specification at the time of posting of this Internet-Draft, and is based on a proposal described in [RFC7942]. The description of implementations in this section is intended to assist the IETF in its decision processes in progressing drafts to RFCs. Please note that the listing of any individual implementation here does not imply endorsement by the IETF. Furthermore, no effort has been spent to verify the information presented here that was supplied by IETF contributors. This is not intended as, and must not be construed to be, a catalog of available implementations or their features. Readers are advised to note that other implementations may exist.

According to [RFC7942], "this will allow reviewers and working groups to assign due consideration to documents that have the benefit of running code, which may serve as evidence of valuable experimentation and feedback that have made the implemented protocols more mature. It is up to the individual working groups to use this information as they see fit".

5.1. Huawei Technologies

*Organization: Huawei Technologies.

*Implementation: Huawei PTN7900 Series Routers implementation of SR-TP[HW-IMP].
*Description: SR-TP is a feature of Huawei PTN7900 series Routers, which uses Path Segments to associate with paths and build up bidirectional paths. Huawei PTN7900 Series Routers with version V100R018C00 and above have commercially implemented the definition of Path Segment and use cases which is defined in section 2 and Section 3.2 in this document, including all the "MUST" and "SHOULD" clauses, while other use cases for Path Segment in section 3 are not yet implemented. For control plane, PTN7900 Series Routers support configuring Path Segment using NETCONF.

*Maturity Level: Product

*Coverage: Partial, section 2 and use case section 3.2.

*Version: Draft-12

*Licensing: N/A

*Implementation experience: Nothing specific.

*Contact: li.fan@huawei.com

*Last updated: September 14, 2023

5.2. ZTE Corp

*Organization: ZTE Corporation.

*Implementation: ZTE's SPN implementation of path segment[ZTE-IMP].

*Description: The feature of SR-MPLS Path Segment has been implemented in ZTE SPN products and follows the definition and mechanism as defined in section 2 and Section 3.2 including all the "MUST" and "SHOULD" clauses while other use cases for Path Segment in section 3 are not yet implemented.

*Maturity Level: Product

*Coverage: Partial, section 2 and use case section 3.2.

*Version: Draft-12

*Licensing: N/A

*Implementation experience: Nothing specific.

*Contact: liu.aihua@zte.com.cn
5.3. **New H3C Technologies**


*Description: Section 2 and Section 3.2 including all the "MUST" and "SHOULD" clauses have been implemented in above mentioned New H3C Products (running Version 7.1.086 and above) for testing, while other use cases for Path Segment in section 3 are not yet implemented.

*Maturity Level: Beta

*Coverage: Partial, section 2 and use case section 3.2.

*Version: Draft-12

*Licensing: N/A

*Implementation experience: Nothing specific.

*Contact: linchangwang.04414@h3c.com

*Last updated: September 13, 2023

5.4. **Spirent Communications**

*Organization: Spirent Communications

*Implementation: Spirent Testcenter Product Family implementation of SR-TP test capability [SP-IMP].

*Description: Spirent Testcenter product family implements SR-MPLS path segment test capabilities on the versions above Spirent Testcenter 4.85. Spirent Testcenter fully support testing all clauses defined in section 2 and Section 3.1,3.2,3.4 , including all the "MUST" and "SHOULD" clauses, and partially support the test of clauses in section 3.3.

*Maturity Level: Production

*Coverage: fully cover section 2 and use case section 3.1,3.2,3.4, partially cover section 3.3

*Version: Draft-12
5.5. Fiberhome

*Organization: Fiberhome Corporation.

*Implementation: Fiberhome SPN series of products (Citrans 650/690E) implementation of path segment [FH-IMP].

*Description: SR-TP is a feature of SPN products which realizes a controllable L3 tunnel, builds the end-to-end L3 deployment business model. The path segment follows the definition and mechanism as defined in section 2 and Section 3.2 including all the "MUST" and "SHOULD" clauses had been implemented, while other use cases for Path Segment in section 3 are not yet implemented.

*Maturity Level: Product

*Coverage: Partial, section 2 and use case section 3.2.

*Version: Draft-12

*Licensing: N/A

*Implementation experience: Nothing specific.

*Contact: zhhan@fiberhome.com

*Last updated: September 21, 2023

5.6. Interoperability Test

[Note to the RFC Editor - remove this section before publication, as well as remove the reference to [RFC7942].]

The Interoperability test of path segment had been done among products from several vendors, including Huawei(PTN7900, V100R018C00), ZTE(ZXCTN 6180, Ver 4.00.00), FiberHome(Citrans 650/690E), Spirent (Chassis: SPT-N4U-220.Test. Module: PX3-QSFP28-12-225A. Version: 4.86) and Nokia in 2018[INTEROP-TEST]. Note that Path Segment is a key feature of Layer3 in SPN architecture [SPN-L3]. This is reported by Weiqiang Cheng from China Mobile at September, 21, 2023.
6. IANA Considerations

This document does not require any IANA actions.

7. References

7.1. Normative References


7.2. Informative References


Acknowledgements

The authors would like to thank Adrian Farrel, Stewart Bryant, Shuangping Zhan, Alexander Vainshtein, Andrew G. Malis, Ketan Talaulikar, Shraddha Hegde, Xinyue Zhang, Loa Andersson and Bruno Decraene for their review, suggestions, comments and contributions to this document.

The authors would like to acknowledge the contribution from Alexander Vainshtein on "Nesting of Path Segments".

Contributors

The following people have substantially contributed to this document.

Mach (Guoyi) Chen
Huawei Technologies Co., Ltd

Email: mach.chen@huawei.com

Lei Wang
China Mobile

Email: wangleiyj@chinamobile.com

Aihua Liu
ZTE Corp

Email: liu.aihua@zte.com.cn
Greg Mirsky
ZTE Corp

Email: gregimirsky@gmail.com

Gyan S. Mishra
Verizon Inc.

Email: gyan.s.mishra@verizon.com

Authors' Addresses

Weiqiang Cheng (editor)
China Mobile

Email: chengweiqiang@chinamobile.com

Han Li
China Mobile

Email: lihan@chinamobile.com

Cheng Li (editor)
Huawei Technologies Co., Ltd
China

Email: c.l@huawei.com

Rakesh Gandhi
Cisco Systems, Inc.
Canada

Email: rgandhi@cisco.com

Roya Zigler
Broadcom

Email: royi.zigler@broadcom.com