SPRING Working Group
Internet-Draft
Intended status: Standards Track
Expires: April 19, 2019

W. Cheng
L. Wang
H. Li
China Mobile
M. Chen
Huawei
R. Gandhi
Cisco Systems, Inc.
R. Zigler
Broadcom
S. Zhan
ZTE

October 16, 2018

Path Segment in MPLS Based Segment Routing Network
draft-cheng-spring-mpls-path-segment-03

Abstract

A Segment Routing (SR) path is identified by an SR segment list, one or partial segments of the list cannot uniquely identify the SR path. Path identification is a pre-requisite for various use-cases such as performance measurement (PM) of an SR path.

This document defines a new type of segment that is referred to as Path Segment, which is used to identify an SR path. When used, it is inserted at the ingress node of the SR path and immediately follows the last segment of the SR path. The Path Segment will not be popped off until it reaches the egress node of the SR path.

Path Segment can be used by the egress node to implement path identification hence to support various use-cases including SR path PM, end-to-end 1+1 SR path protection and bidirectional SR paths correlation.

Status of This Memo

Table of Contents

1. Introduction

Segment Routing (SR) [RFC8402] is a source routed forwarding method that allows to directly encode forwarding instructions (called segments) in each packet, hence it enables to steer traffic through a network without the per-flow states maintained on the transit nodes. Segment Routing can be instantiated on MPLS data plane or IPv6 data plane. The former is called SR-MPLS, the latter is called SRv6

[RFC8402]. SR-MPLS leverages the MPLS label stack to construct SR path, and SRv6 uses the a new IPv6 Extension Header (EH) called the IPv6 Segment Routing Header (SRH) [I-D.ietf-6man-segment-routing-header] to construct SR path.

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. So 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 determine from which SR path the packet comes.

However, to support use cases like end-to-end 1+1 path protection (Live-Live case), bidirectional path correlation or performance measurement (PM), the ability to implement path identification is a pre-requisite.

Therefore, this document introduces a new segment that is referred to as Path Segment. A Path Segment is defined to uniquely identify an SR path in the context of the egress node. It is normally used by egress nodes for path identification or correlation.

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 BCP14 [RFC2119][RFC8174] when, and only when, they appear in all capitals, as shown here.

1.2. Abbreviations

DM: Delay Measurement.

LM: Loss Measurement.

MPLS: Multiprotocol Label Switching.

PM: Performance Measurement.

PSID: Path Segment ID.

SID: Segment ID.

SL: Segment List.

SR: Segment Routing.

SR-MPLS: Segment Routing instantiated on MPLS data plane.

SRv6: Segment Routing instantiated on IPv6 data plane

2. Path Segment

A Path Segment is a single label that is assigned from the Segment Routing Local Block (SRLB) or Segment Routing Global Block (SRGB) of the egress node of an SR path. It means that the Path Segment is unique in the context of the egress node of the SR path. When Path Segment is used, the Path Segment MUST be inserted at the ingress node and MUST immediately follow the last label of the SR path. The Path Segment may be used to identify an SR-MPLS Policy, its Candidate-Path (CP) or a SID List (SL) [I-D.ietf-spring-segment-routing-policy] terminating on an egress node depending on the use-case.

The value of the TTL field of the Path Segment MUST be set to the same value of the last segment label of the SR path. If the Path Segment is the bottom label, the S bit MUST be set.

Normally, the intermediate nodes will not see the Path Segment label and do not know how to process it. A Path Segment presenting to an intermediate node is an error condition.

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

The label stack with Path Segment is as below (Figure1):

```
            +--------------------+
            |       ...          |
            +--------------------+
            |      Label 1       |
            +--------------------+
            |      Label 2       |
            +--------------------+
            |       ...          |
            +--------------------+
            |      Label n       |
            +--------------------+
            |     Path Segment   |
            +--------------------+
            |       ...          |
            +--------------------+
```

Figure 1: Label Stack with Path Segment

Where:

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

- o The Path Segment identifies the SR path in the context of the egress node of the SR path.

3. Nesting of Path Segments

Binding SID (BSID) [RFC8402] can be used for SID list compression. With BSID, an end-to-end SR path 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 Path SID (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) that spans three 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 allocated a BSID. For nesting the sub-paths, each sub-path is allocated a PSID. Then, 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.

```
            /--------\       /--------\       /--------\
          /            \   /            \   /            \
        A{    Access    }B{  Aggregation }C{     Core     }D
          \            /   \            /   \            /
            \--------/       \--------/       \--------/
          Sub-path(A->B)    Sub-path(B->C)   Sub-path(C->D)
       |<--------------->|<-------------->|<-------------->|
                             E2E Path(A->D)
       |<------------------------------------------------->|

    +------------+
    ~A->B SubPath~
    +------------+  +------------+
    |s-PSID(A->B)|  ~B->C SubPath~
    +------------+  +------------+
    | BSID(B->C) |  |s-PSID(B->C)|
    +------------+  +------------+  +------------+
    | BSID(C->D) |  | BSID(C->D) |  ~C->D SubPath~
    +------------+  +------------+  +------------+  +------------+
    |e-PSID(A->D)|  |e-PSID(A->D)|  |e-PSID(A->D)|  |e-PSID(A->D)|
    +------------+  +------------+  +------------+  +------------+
```

Figure 2: Nesting of Path Segments

4. Path Segment Allocation

Several ways can be used to allocate the Path Segment.

One way is to set up a communication channel (e.g., MPLS Generic Associated Channel (G-ACh)) between the ingress node and the egress node, and the ingress node of the SR path can directly send a request to the egress node to ask for a Path Segment.

Another way is to leverage a centralized controller (e.g., PCE, SDN controller) to assign the Path Segment. PCEP based Path Segment allocation is defined in [I-D.li-pce-sr-path-segment], and SR-policy based path segment allocation is defined in [I-D.li-idr-sr-policy-path-segment-distribution].

5. Path Segment for PM

As defined in [RFC7799], performance measurement can be classified into Active, Passive and Hybrid measurement. For Passive measurement, path identification at the measuring points is the pre-requisite. Path segment can be used by the measuring points (e.g., the ingress/egress nodes of an SR path) or a centralized controller to correlate the packets counts/timestamps that are from the ingress

and egress nodes to a specific SR path, then packet loss/delay can be calculated.

Performance Delay Measurement (DM) and Loss Measurements (LM) in SR networks with MPLS data plane can be found in [I-D.gandhi-spring-sr-mpls-pm] and [I-D.gandhi-spring-udp-pm].

6. Path Segment for Bi-directional SR Path

With the current SR architecture, an SR path is a unidirectional path. In some scenarios, for example, mobile backhaul transport network, there are requirements to support bidirectional path, and the path is normally treated as a single entity and both directions of the path have the same fate, for example, failure in one direction will result in switching at both directions.

MPLS supports this by introducing the concepts of co-routed bidirectional LSP and associated bidirectional LSP. With SR, to support bidirectional path, a straightforward way is to bind two unidirectional SR paths to a single bidirectional path. Path segments can be used to correlate the two unidirectional SR paths at both ends of the paths.

[I-D.li-pce-sr-bidir-path] defines how to use PCEP and Path segment to initiate a bidirectional SR path, and [I-D.li-idr-sr-policy-path-segment-distribution] defines how to use SR policy and Path segment to initiate a bidirectional SR path.

7. 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 an SR path needs to know the set of paths that constitute the primary and the secondary(s), in order to select the primary packet for onward transmission, and to discard the packets from the secondary(s).

To do this, each path needs a path identifier that is unique at the egress node. Depending on the design, this is a single unique path segment label chosen by the egress PE.

There then needs to be a method of binding this path identifiers into equivalence groups such that the egress PE can determine the set of packets that represent a single path and its secondary.

It is obvious that this group can be instantiated in the network by an SDN controller.

8. IANA Considerations

This document does not require any IANA actions.

9. Security Considerations

This document does not introduce additional security requirements and mechanisms other than the ones described in [RFC8402].

10. Contributors

The following individuals also contribute to this document.

- Cheng Li, Huawei

11. Acknowledgements

The authors would like to thank Stewart Bryant, Alexander Vainshtein, Andrew G. Malis and Loa Andersson for their review, suggestions and comments to this document.

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

12. References

12.1. Normative References

[I-D.ietf-spring-segment-routing-mpls]
Bashandy, A., Filsfils, C., Previdi, S., Decraene, B., Litkowski, S., and R. Shakir, "Segment Routing with MPLS data plane", draft-ietf-spring-segment-routing-mpls-14 (work in progress), June 2018.

[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>.

[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>.

12.2. Informative References

[I-D.gandhi-spring-sr-mpls-pm]
Gandhi, R., Filsfils, C., daniel.voyer@bell.ca, d., Salsano, S., Ventre, P., and M. Chen, "Performance Measurement in Segment Routing Networks with MPLS Data Plane", draft-gandhi-spring-sr-mpls-pm-03 (work in progress), September 2018.

[I-D.gandhi-spring-udp-pm]
Gandhi, R., Filsfils, C., daniel.voyer@bell.ca, d., Salsano, S., Ventre, P., and M. Chen, "UDP Path for In-band Performance Measurement for Segment Routing Networks", draft-gandhi-spring-udp-pm-02 (work in progress), September 2018.

[I-D.ietf-6man-segment-routing-header]
Filsfils, C., Previdi, S., Leddy, J., Matsushima, S., and d. daniel.voyer@bell.ca, "IPv6 Segment Routing Header (SRH)", draft-ietf-6man-segment-routing-header-14 (work in progress), June 2018.

[I-D.ietf-spring-segment-routing-policy]
Filsfils, C., Sivabalan, S., daniel.voyer@bell.ca, d., bogdanov@google.com, b., and P. Mattes, "Segment Routing Policy Architecture", draft-ietf-spring-segment-routing-policy-01 (work in progress), June 2018.

[I-D.li-idr-sr-policy-path-segment-distribution]
Li, C., Chen, M., Dong, J., and Z. Li, "Segment Routing Policies for Path Segment and Bi-directional Path", draft-li-idr-sr-policy-path-segment-distribution-00 (work in progress), April 2018.

[I-D.li-pce-sr-bidir-path]
Li, C., Chen, M., Dhody, D., Cheng, W., Li, Z., Dong, J., and R. Gandhi, "PCEP Extension for Segment Routing (SR) Bi-directional Associated Paths", draft-li-pce-sr-bidir-path-01 (work in progress), September 2018.

[I-D.li-pce-sr-path-segment]
Li, C., Chen, M., Dhody, D., Cheng, W., Dong, J., Li, Z., and R. Gandhi, "Path Computation Element Communication Protocol (PCEP) Extension for Path Identification in Segment Routing (SR)", draft-li-pce-sr-path-segment-02 (work in progress), September 2018.

[RFC6374] Frost, D. and S. Bryant, "Packet Loss and Delay Measurement for MPLS Networks", RFC 6374, DOI 10.17487/RFC6374, September 2011, <https://www.rfc-editor.org/info/rfc6374>.

[RFC7799] Morton, A., "Active and Passive Metrics and Methods (with Hybrid Types In-Between)", RFC 7799, DOI 10.17487/RFC7799, May 2016, <https://www.rfc-editor.org/info/rfc7799>.

[RFC8321] Fioccola, G., Ed., Capello, A., Cociglio, M., Castaldelli, L., Chen, M., Zheng, L., Mirsky, G., and T. Mizrahi, "Alternate-Marking Method for Passive and Hybrid Performance Monitoring", RFC 8321, DOI 10.17487/RFC8321, January 2018, <https://www.rfc-editor.org/info/rfc8321>.

Authors' Addresses

Weiqiang Cheng
China Mobile

Email: chengweiqiang@chinamobile.com

Lei Wang
China Mobile

Email: wangleiyj@chinamobile.com

Han Li
China Mobile

Email: lihan@chinamobile.com

Mach(Guoyi) Chen
Huawei

Email: mach.chen@huawei.com

Rakesh Gandhi
Cisco Systems, Inc.
Canada

Email: rgandhi@cisco.com

Royi Zigler
Broadcom

Email: royi.zigler@broadcom.com

Shuangping Zhan
ZTE

Email: zhan.shuangping@zte.com.cn