SPRING Working Group   W. Cheng
Internet-Draft   H. Li
Intended status: Standards Track   China Mobile
Expires: March 27, 2021   M. Chen
                                                             Huawei
                                                          R. Gandhi
                                              Cisco Systems, Inc.
                                                          R. Zigler
                                                           Broadcom
                                                September 23, 2020

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

Abstract

   A Segment Routing (SR) path is identified by an SR segment list.
   Only the complete segment list can identify the end-to-end SR path,
   and 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), bidirectional paths correlation, and
   end-to-end 1+1 path protection.

   In SR for MPLS data plane (SR-MPLS), the segment identifiers are
   stripped from the packet through label popping as the packet transits
   the network. This means that when a packet reaches the egress of the
   SR path, it is not possible to determine on which SR path it
   traversed the network.

   This document defines a new type of segment that is referred to as
   Path Segment, which is used to identify an SR path in an SR-MPLS
   network. When used, it is inserted by the ingress node of the SR
   path and immediately follows the last segment identifier in the
   segment list of the SR path. The Path Segment will not be popped off
   until it reaches the egress node of the SR path. The Path Segment
   then can be used by the egress node to implement SR path
   identification and correlation.

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 March 27, 2021.

Copyright Notice

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 steering traffic through a
network without the per-flow states maintained on the transit nodes.
Segment Routing can be instantiated on an MPLS data plane or an IPv6
data plane. The former is called SR-MPLS [RFC8660], the latter is
called SRv6 [RFC8402]. SR-MPLS leverages the MPLS label stack to
construct an 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 (e.g. Explicit-Null label) may
be left in the MPLS label stack when the packet reaches the egress
node. Thus, the egress node cannot determine along which SR path the
packet came.

However, to support various use-cases in SR-MPLS networks, like end-
to-end 1+1 path protection (Live-Live case) [RFC4426], bidirectional
path [RFC5654], or Performance Measurement (PM) [RFC7799], the
ability to implement path identification on the egress node is a pre-
requisite.

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

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
[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.

MSD: Maximum SID Depth.

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.

[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) or
dynamic MPLS label pool 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 a 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, in other words, inserted after the routing
segment (adjacency/node/prefix segment) pointing to the egress node.

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 in the MPLS label stack entry containing
the Path Segment MUST be set to the same value as the TTL of the last
label stack entry for the last segment in 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.

A Path Segment can be used in the case of Penultimate Hop Popping
(PHP), where some labels are be popped off at the penultimate hop of
an SR path, but the Path Segment MUST NOT be popped off until it
reaches at the egress node.

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.

In some deployments, service labels may be added after the Path
Segment label in the MPLS label stack. In this case, the egress node
MUST be capable of processing more than one label. The additional
processing required, may have an impact on forwarding performance.

Generic Associated Label (GAL) is used for Operations, Administration
and Maintenance (OAM) in MPLS networks [RFC5586]. When GAL is used,
it MUST be added at the bottom of the label stack after the Path
Segment label.

Entropy label and Entropy Label Indicator (ELI) as described in
[RFC8662] for SR-MPLS path, can be placed before or after the Path
Segment label in the MPLS label stack.

The SR path computation needs to know the Maximum SID Depth (MSD)
that can be imposed at each node/link 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. The MSD
used for path computation MUST include the Path Segment label.

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

```
            +--------------------+
            |        ...         |
            +--------------------+
            |      Label 1       |
            +--------------------+
            |      Label 2       |
            +--------------------+
            |        ...         |
            +--------------------+
            |      Label n       |
            +--------------------+
            |    Path Segment    |
            +--------------------+
            |        ...         |
            +--------------------+
            ~       Payload      ~
            +--------------------+
```
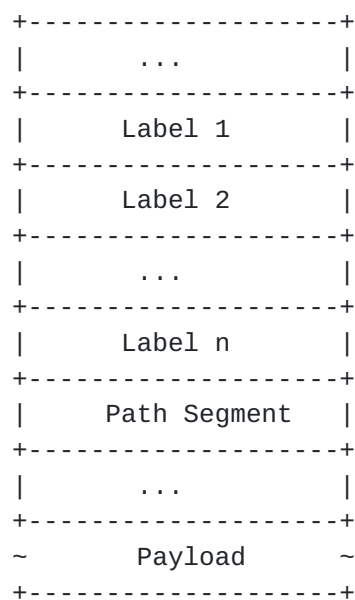
        Figure 1: Label Stack with Path Segment

    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. 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)) [RFC5586] 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 allocate a Path Segment. The
   detail of G-ACh based solution is left for future study and out of
   the scope of this document.

   Another way is to leverage a centralized controller (e.g., SDN
   controller) to assign the Path Segment. In this case, the controller
   will deliver the Path Segment and corresponding path information
   (e.g., SR policy) to the ingress node. Path Computation Element
   Protocol (PCEP) based Path Segment allocation for SR Policy is
   defined in [I-D.ietf-pce-sr-path-segment], BGP based Path Segment
   allocation for SR Policy is defined in
   [I-D.ietf-idr-sr-policy-path-segment]. At the same time, the
   controller MUST make sure (e.g., by some capability discovery
   mechanisms outside the scope of this document) that the egress node
   knows the Path Segment and it can process it, as well as the label
   does not collide with any label allocation done by the egress node.

4. 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
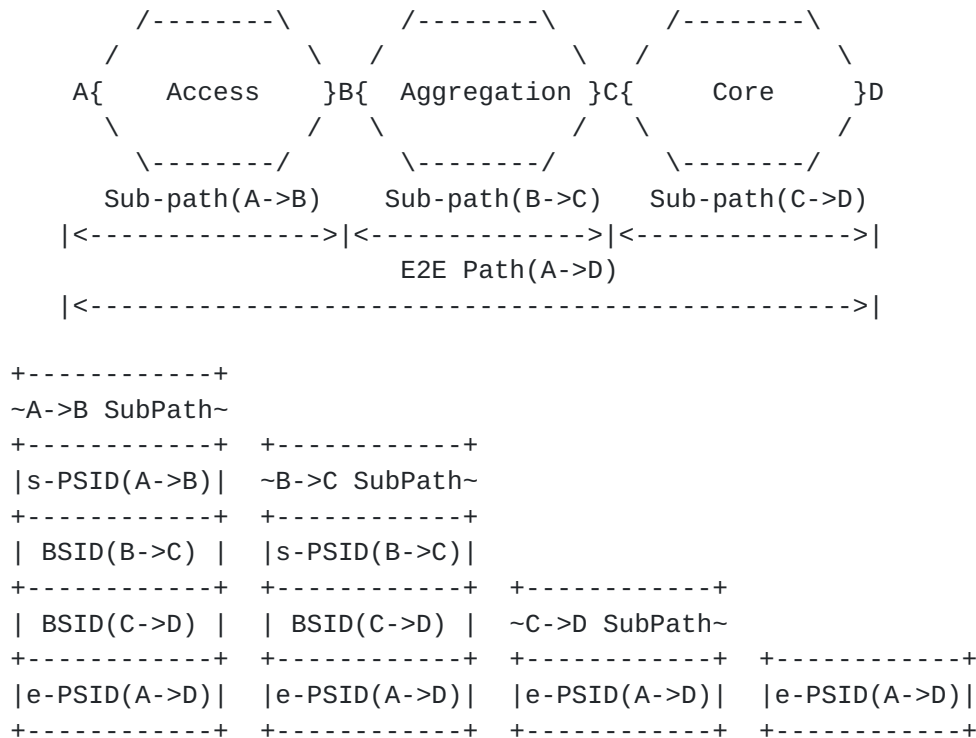multi-domain scenario.

```
       /--------\        /--------\        /--------\
      /          \  /              \  /              \
    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

5. Path Segment for Performance Measurement

   As defined in [RFC7799], performance measurement can be classified
   into Passive, Active, and Hybrid measurement.

   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
   [I-D.ietf-mpls-rfc6374-sr] and [I-D.gandhi-spring-stamp-srpm].

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 [I-D.gandhi-mpls-ioam-sr].

Path Segment can also be used for In-band PM for SR-MPLS to identify the SR Path associated with the collected performance metrics [I-D.cheng-mpls-inband-pm-encapsulation].

6. 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. The both 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.

[I-D.ietf-pce-sr-bidir-path] defines procedures on how to use PCEP for SR Policy to initiate a bidirectional SR path. Also, [I-D.ietf-idr-sr-policy-path-segment] defines procedures on how to use BGP for SR Policy 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 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. It is obvious that this equivalence group can be instantiated in the network by an SDN controller using the Path Segments of the SR paths.

8. Security Considerations

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

9. IANA Considerations

This document does not require any IANA actions.

10. References

10.1. Normative References

[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>.

[RFC8660]   Bashandy, A., Ed., Filsfils, C., Ed., Previdi, S.,
            Decraene, B., Litkowski, S., and R. Shakir, "Segment
            Routing with the MPLS Data Plane", RFC 8660,
            DOI 10.17487/RFC8660, December 2019,
            <https://www.rfc-editor.org/info/rfc8660>.

10.2. Informative References

[I-D.cheng-mpls-inband-pm-encapsulation]
            Cheng, W., Min, X., Zhou, T., Dong, X., and Y. Peleg,
            "Encapsulation For MPLS Performance Measurement with
            Alternate Marking Method", draft-cheng-mpls-inband-pm-
            encapsulation-04 (work in progress), September 2020.

[I-D.gandhi-mpls-ioam-sr]
            Gandhi, R., Ali, Z., Filsfils, C., Brockners, F., Wen, B.,
            and V. Kozak, "MPLS Data Plane Encapsulation for In-situ
            OAM Data", draft-gandhi-mpls-ioam-sr-03 (work in
            progress), September 2020.

   [I-D.gandhi-spring-stamp-srpm]
             Gandhi, R., Filsfils, C., Voyer, D., Chen, M., and B.
             Janssens, "Performance Measurement Using Simple TWAMP
             (STAMP) for Segment Routing Networks", draft-gandhi-
             spring-stamp-srpm-02 (work in progress), August 2020.

   [I-D.ietf-idr-sr-policy-path-segment]
             Li, C., Li, Z., Telecom, C., Cheng, W., and K. Talaulikar,
             "SR Policy Extensions for Path Segment and Bidirectional
             Path", draft-ietf-idr-sr-policy-path-segment-01 (work in
             progress), August 2020.

   [I-D.ietf-mpls-rfc6374-sr]
             Gandhi, R., Filsfils, C., Voyer, D., Salsano, S., and M.
             Chen, "Performance Measurement Using RFC 6374 for Segment
             Routing Networks with MPLS Data Plane", draft-ietf-mpls-
             rfc6374-sr-00 (work in progress), July 2020.

   [I-D.ietf-pce-sr-bidir-path]
             Li, C., Chen, M., Cheng, W., Gandhi, R., and Q. Xiong,
             "PCEP Extensions for Associated Bidirectional Segment
             Routing (SR) Paths", draft-ietf-pce-sr-bidir-path-03 (work
             in progress), September 2020.

   [I-D.ietf-pce-sr-path-segment]
             Li, C., Chen, M., Cheng, W., Gandhi, R., and Q. Xiong,
             "Path Computation Element Communication Protocol (PCEP)
             Extension for Path Segment in Segment Routing (SR)",
             draft-ietf-pce-sr-path-segment-01 (work in progress), May
             2020.

   [I-D.ietf-spring-segment-routing-policy]
             Filsfils, C., Talaulikar, K., Voyer, D., Bogdanov, A., and
             P. Mattes, "Segment Routing Policy Architecture", draft-
             ietf-spring-segment-routing-policy-08 (work in progress),
             July 2020.

   [RFC4426] Lang, J., Ed., Rajagopalan, B., Ed., and D. Papadimitriou,
             Ed., "Generalized Multi-Protocol Label Switching (GMPLS)
             Recovery Functional Specification", RFC 4426,
             DOI 10.17487/RFC4426, March 2006,
             <https://www.rfc-editor.org/info/rfc4426>.

   [RFC5586] Bocci, M., Ed., Vigoureux, M., Ed., and S. Bryant, Ed.,
             "MPLS Generic Associated Channel", RFC 5586,
             DOI 10.17487/RFC5586, June 2009,
             <https://www.rfc-editor.org/info/rfc5586>.

   [RFC5654] Niven-Jenkins, B., Ed., Brungard, D., Ed., Betts, M., Ed.,
              Sprecher, N., and S. Ueno, "Requirements of an MPLS
              Transport Profile", RFC 5654, DOI 10.17487/RFC5654,
              September 2009, <https://www.rfc-editor.org/info/rfc5654>.

   [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>.

   [RFC8662] Kini, S., Kompella, K., Sivabalan, S., Litkowski, S.,
              Shakir, R., and J. Tantsura, "Entropy Label for Source
              Packet Routing in Networking (SPRING) Tunnels", RFC 8662,
              DOI 10.17487/RFC8662, December 2019,
              <https://www.rfc-editor.org/info/rfc8662>.

   [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>.

Acknowledgements

Contributors

   The following people have substantially contributed to this document:

      Cheng Li
      Huawei Technologies

      EMail: chengli13@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

Authors' Addresses

   Weiqiang Cheng
   China Mobile

   Email: chengweiqiang@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