A Simplified Scalable E-Line Service Model with Segment Routing Underlay
                draft-boutros-bess-eline-services-over-sr-00

Abstract

   This document proposes a new approach for realizing Ethernet line
   (E-Line) services over Segment Routing (SR) networks. This approach
   significantly improves scalability and convergence of control plane,
   and simplifies network operation. Furthermore, it naturally yields
   All-Active multi-homing support for E-Line services without relying
   on any overlay techniques.

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 May 6, 2021.

Copyright Notice

Table of Contents

1. Introduction

Historically, E-Line services are realized as Virtual Private Wire
Services (VPWS) using point-to-point (P2P) Pseudo-Wires (PWs). A PW
identifies both the service type and the service termination nodes in
both control and data planes. A PW can be dynamically established
via LDP or BGP. Ethernet VPN (EVPN) signaling mechanisms via BGP can
be used to provide VPWS with Single-Active and All-Active multihoming
redundancy as well as Inter-Autonomous System (AS) support.

This document proposes a new approach for supporting E-Line services
over Segment Routing (SR) networks. It reduces BGP signaling
overhead by at least by two orders of magnitudes compared to the
current EVPN-VPWS signaling mechanisms and hence yields better

control plane scalabilty. Furthermore, it eliminates the need to
associate Route Distinguisher (RD) and Route Target (RT) with EVPN
routes, and hence leads to simpler network operations. The proposed
approach enables the benefit of All-Active redundancy and aliasing
for the Multi-Home (MH) sites. This proposal makes use of the
properties of SR anycast SID to achieve redundancy and fast
convergence at the transport level. Anycast SIDs provide the ability
to discover nodes supporting multihoming and perform Designated
Forwarder (DF) election. As such, the need for any overlay mechanism
to achieve redundancy and fast convergence is eliminated. Also, the
proposed approach supports auto-discovery and single-sided service
provisioning.

In the proposed approach, an E-Line service instance is represented
by a Segment ID (SID) which is referred to "Service SID" in the rest
of the document. Such a SID can be:

o an MPLS label for SR-MPLS.

o a uSID (micro SID) for SRv6 representing network function
   associated with an E-Line service instance. A new SRv6 network
   programming function will be specified in the next version of this
   document.

In the present form, classical VPWS service is incapable of
supporting All-Active multihoming. However, thanks to SR anycast SID
capability, the proposed approach natively provides such redundancy
at transport layer.

A Service SID is unique within a service domain. A node can
advertise service SID(s) of the E-Line instance(s) that it hosts via
BGP for auto-discovery purpose. In the case of SR-MPLS, a service
SID can be carried as a range of absolute MPLS label values or an
index into an Segment Routing Global Block (SRGB), and in the case of
SRv6, a service SID can be carried as uSID in BGP update.

A node advertising E-Line service routes packs information about as
many service instances as possible at the time of advertisement in a
single route. The objective is to reduce the volume of signaling
messages advertised as well as processed in control plane. E-Line
service SIDs are represented as an array. A service route contains
SID value at the start of the array and a bitmask where each bit
represents an index in the array. The SID value for an E-Line
service instance can be derived from the start and the index of the
array. When advertising routes to E-Line services, a node sends the
initial value of the SID array and sets the bitmask to indicate all
E-Line services instances that it hosts along with its node SID which
may be anycast SID for MH case. For FXC, a route containing the

start normalized VID and the bitmask of VIDs in association with the E-Line service SID is advertised. The necessary BGP extension will be described in a future version of this document.

Each node attached to the MH site advertises the same anycast SID to allow other nodes to discover the redundancy group membership (auto-discovery).

The proposed solution can also be applicable to the existing EVPN control plane without compromising its benefits such as All-Active multihoming on access, aliasing in the core, auto-provisioning and auto-discovery, etc. With this approach, the need for advertising of EVPN route types 1 and 4 as well Split-Horizon (SH) label is eliminated which simplifies EVPN-VPWS operations.

In the following sections, we will describe the functionalities of the proposed approach in detail.

[2](). Terminology

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [[RFC2119]()].

[3](). Abbreviations

CE: Customer Edge node e.g., host or router or switch.

E-Line: Ethernet virtual private line.

EVPN: Ethernet VPN.

MH: Multi-Home.

OAM: Operations, Administration and Maintenance.

PE: Provide Edge Node.

SID: Segment Identifier.

SR: Segment Routing.

VPWS: Virtual Private Wire Service.

4. Control Plane Behavior

```
                            ____ CE3
                        /                    ____CE1
                 -------- PE3 --------  /
                 /                      PE1
                /                       | \
              PE5                       |  \
             /|                         |   \ CE2
            / | Service Provider Network |   /
        CE5  |                          |  /
            \ |                         | /
             \|                         PE2
            PE6                        /
            /  -------- PE4  --------
       CE6___ /     CE4_____/
```
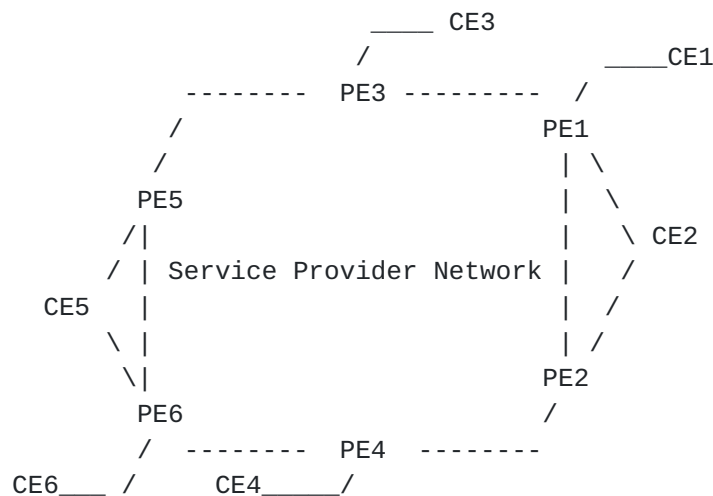
        Figure 1: A reference network used for the examples below

4.1. Service discovery

   A node (e.g., PE5 in Figure 1) can discover E-Line instances as well
   as the associated service SIDs on other nodes (e.g., PE1, PE2, etc in
   Figure 1) via configuration or auto-discovery. With the latter, the
   service SIDs may be advertised using BGP.

4.2. All-Active Service Redundancy

   An anycast SID per Ethernet Segment (ES) will be configured on all
   nodes attached to a Multi-Home (MH) site. For example, in Figure 1,
   PE1 and PE2 are configured with an anycast SID representing the
   multi-homed site CE2. The ES anycast SIDs will be advertised in BGP
   by nodes (e.g., PE1 and PE2) connected to the MH site.

   Upon receiving advertisement containing ES anycast SID, a node (e.g.,
   PE5 in Figure 1) learns the nodes (e.g., PE1 and PE2) hosting MH
   sites, and performs DF election. Aliasing is natively achieved at
   transport layer.

4.3. E-Line Attributes

   Layer 2 extended community can be advertised with E-Line service
   routes. Such a route includes the start of service SID and the
   bitmap of all other Services enabled on the advertising node. This
   will be elaborated more in the next revision of this document. As
   mentioned earlier, the goal is to pack information about as many

E-Line services hosted on the advertising node as possible to reduce the overall amount of signaling messages.

4.4. Service withdrawal

Node failure can be detected due via IGP convergence. For faster detection of node failure, mechanism like BFD can be deployed. The proposed approach does not require additional service withdrawal mechanism.

On PE-CE link failure, the PE node withdraws the route to the corresponding ES in BGP in order to stop receiving traffic to that ES.

With MH case with anycast SID, upon detecting a failure on PE-CE link, a PE node may forward incoming traffic to the impacted ES(s) to another PE node which is part of the anycast group until it withdraws routes to the impacted ES(s) for faster convergence. For example, in Figure 1, assuming PE5 and PE6 are part of an anycast group, upon link failure between PE5 and CE5, PE5 can forward the received packets from the core to PE6 until it withdraws the anycast SID associated with the MH site.

5. Data Plane Behavior

5.1. Data Packet Format

The proposed method requires unicast data packet be formed as shown in Figure 2.

```
            +-------------------------------+
            | SID(s) to reach destination   |
            +-------------------------------+
            |           Service SID         |
            +-------------------------------+
            |         Layer-2 Payload       |
            +-------------------------------+
```

Figure 2: Data packet format for sending traffic to a destination

o SID(s) to reach destination: depends on the intent of the underlay transport:

   * IGP shortest path: node SID of the destination. The destination can belong to an anycast group.

* IGP path with intent: Flex-Algo SID if the destination can be reached using the Flex-Algo SID for a specific intent (e.g., low latency path). The destination can belong to an anycast group.

* SR policy: a SID-list for the SR policy that can be used to reach the destination.

o Service SID: The SID that uniquely identifies a E-Line service instance in an SR domain.

5.2. Aliasing

Packets destined to a MH CE is distributed to the PE nodes attached to the CE for load-balancing (UCMP/ECMP) purpose. This is achieved implicitly due to the use of anycast SIDs for both ES as well as PE attached to the ES. In our example, traffic destined to CE5 is distributed via PE5 and PE6.

5.3. Single-Homed CE

Referring to Figure 1, PE3 and PE4 provide an E-Line service to connect CE3 to CE4. If PE3 wants to forward a packet received from CE3 to CE4, it formulates the packet as shown in Figure 3.

```
          +-----------------------------+
          |Transport SID(s) to reach PE4|
          +-----------------------------+
          |          Service SID        |
          +-----------------------------+
          |        Layer-2 Packet       |
          +-----------------------------+
```

Figure 3: Data packet format for forwarding traffic from PE3 to CE4

5.4. Multi-Homed CE

Referring to Figure 1, PE5 and PE6 and PE1 and PE2 provide to provide E-Line services to connect CE2 and CE5. PE5 and PE6 share an anycast SID, and PE1 and PE2 share another anycast SID.

Packets sent from CE2 to CE5 are load-balanced (ECMP/UCMP) between PE5 and PE6 assuming that PE1 and PE2 learned the corresponding E-Line Service SID from PE5 and PE6.

The following diagram shows SID stack for a L2 packet sent from CE2 to CE5.

```
+--------------------------------+
|Anycast Node SID for PE5 and PE6|
+--------------------------------+
|          Service SID           |
+--------------------------------+
|          Layer-2 Packet        |
+--------------------------------+
```

 Figure 4: Data packet format for forwarding traffic from PE1 or PE2 to CE5

   In case of FXC the signaled normalized VID will be encoded in the
   Layer 2 packet.

6. Benefits of E-Line services over SR

   The proposed approach eliminates the need for establishing and
   maintaining PWs as with legacy VPWS technology. This yields
   significant reduction in control plane overhead. The proposed
   approach provides the benefits as such fast convergence. Finally,
   using anycast SID, the proposed approach provides All-Active
   multihoming as well as aliasing.

7. Security Considerations

   The mechanisms in this document use Segment Routing control plane as
   defined in Security considerations described in Segment Routing
   control plane are equally applicable.

8. IANA Considerations

   TBD.

9. Acknowledgements

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>.

   [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>.

   [RFC8754] Filsfils, C., Ed., Dukes, D., Ed., Previdi, S., Leddy, J.,
              Matsushima, S., and D. Voyer, "IPv6 Segment Routing Header
              (SRH)", RFC 8754, DOI 10.17487/RFC8754, March 2020,
              <https://www.rfc-editor.org/info/rfc8754>.

10.2. Informative References

   [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.

   [RFC4664] Andersson, L., Ed. and E. Rosen, Ed., "Framework for Layer
              2 Virtual Private Networks (L2VPNs)", RFC 4664,
              DOI 10.17487/RFC4664, September 2006,
              <https://www.rfc-editor.org/info/rfc4664>.

   [RFC8214] Boutros, S., Sajassi, A., Salam, S., Drake, J., and J.
              Rabadan, "Virtual Private Wire Service Support in Ethernet
              VPN", RFC 8214, DOI 10.17487/RFC8214, August 2017,
              <https://www.rfc-editor.org/info/rfc8214>.

Authors' Addresses

   Sami Boutros (editor)
   Ciena Corporation
   USA

   Email: sboutros@ciena.com

   Siva Sivabalan (editor)
   Ciena Corporation
   Canada

   Email: ssivabal@ciena.com

James Uttaro
AT&T
USA

Email: ju1738@att.com

Daniel Voyer
Bell Canada
Canada

Email: daniel.voyer@bell.ca

Bin Wen
Comcast
USA

Email: bin_wen@cable.comcast.com

Luay Jalil
Verizon
USA

Email: luay.jalil@verizon.com