INTERNET-DRAFT
Intended Status: Standard Track
Sami Boutros
Ali Sajassi
Samer Salam
Cisco Systems
John Drake
Juniper Networks
Jeff Tantsura
Ericsson
Expires: August 28, 2013
February 24, 2013

VPWS support in E-VPN
draft-boutros-l2vpn-evpn-vpws-01.txt

Abstract

This document describes how E-VPN can be used to support virtual private wire service (VPWS) in MPLS/IP networks. E-VPN enables the following characteristics for VPWS: active/standby as well as active/active multi-homing with flow-based load-balancing, eliminates the need for single-segment and multi-segment PW signaling, and provides fast protection using data-plane prefix independent convergence upon node or link failure.

Status of this Memo

Table of Contents

1 Introduction

This document describes how E-VPN can be used to support virtual private wire service (VPWS) in MPLS/IP networks. The use of E-VPN mechanisms for VPWS applies the benefits of E-VPN to p2p services. These benefits include active/standby AC redundancy as well as active/active multi-homing with flow-based load-balancing. Furthermore, the use of E-VPN for VPWS eliminates the need for signaling single-segment and multi-segment PWs for p2p Ethernet services.

[E-VPN] has the ability to forward customer traffic to/from a given customer Attachment Circuit (AC), aka Ethernet Segment in E-VPN terminology, without any MAC lookup. This capability is ideal in providing p2p services (aka VPWS services). [MEF] defines EVPL service as p2p service between a pair of ACs (designated by VLANs). EVPL can be considered as a VPWS with only two ACs. In delivering an EVPL service, the traffic forwarding capability of E-VPN using only a pair of Ethernet AD routes is used; whereas, for more general VPWS, traffic forwarding capability of E-VPN using a group of Ethernet AD routes (one Ethernet AD route per AC/segment) is used. Since in VPWS services, the traffic from an originating Ethernet Segment can go only to a single destination Ethernet Segment, no MAC lookup is needed and the MPLS label associated with the per-EVI Ethernet AD route can be used in forwarding user traffic to the destination AC.

In current PW redundancy mechanisms, convergence time is a function of control plane convergence characteristics. However, with E-VPN it is possible to attain faster convergence through the use of data-plane prefix independent convergence upon node or link failure.

This document proposes the use of the Ethernet AD route to signal labels for P2P Ethernet services. As with E-VPN, the Ethernet Segment route can be used to synchronize state between the PEs attached to the same multi-homed Segment.

1.1 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 RFC 2119 [RFC2119].

MAC: Media Access Control

MPLS: Multi Protocol Label Switching.

OAM: Operations, Administration and Maintenance.

PE: Provide Edge Node.

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

EVI: E-VPN Instance.

2. BGP Extensions

[E-VPN] defines a new BGP NLRI for advertising different route types for E-VPN operation. This document does not define any new BGP messages, but rather re-purposes one of the routes as described next.

This document proposes the use of the per EVI Ethernet AD route to signal P2P services. The Ethernet Segment Identifier field is set to the ESI of the attachment circuit of the VPWS service instance. The Ethernet Tag field is set to 0 in the case of an Ethernet Private Wire service, and to the VLAN identifier associated with the service for Ethernet Virtual Private Wire service. The route is associated with a Route-Target (RT) extended community attribute that identifies the service instance (together with the Ethernet Tag field when non-zero).

3 Operation

The following figure shows an example of a P2P service deployed with E-VPN.

```
          Ethernet                                          Ethernet
          Native   |<---------E-VPN Instance------------>|  Native
          Service  |                                     |  Service
           (AC)    |     |<-PSN1->|       |<-PSN2->|     |  (AC)
             |     V     V        V       V        V     V  |
             |     +-----+      +-----+  +-----+   +-----+  |
      +----+ |     | PE1 |======|ASBR1|==|ASBR2|===| PE3 |  |    +----+
      |    |-------+-----+      +-----+  +-----+   +-----+-------|    |
      | CE1| |                                              |    |CE2 |
      |    |-------+-----+      +-----+  +-----+   +-----+-------|    |
      +----+ |     | PE2 |======|ASBR3|==|ASBR4|===| PE4 |  |    +----+
           ^       +-----+      +-----+  +-----+   +-----+          ^
           |   Provider Edge 1        ^        Provider Edge 2      |
           |                          |                             |
           |                          |                             |
           |              E-VPN Inter-provider point                |
           |                                                        |
           |<---------------- Emulated Service -------------------->|
```

iBGP sessions will be established between PE1, PE2, ASBR1 and ASBR3, possibly via a BGP route-reflector. Similarly, iBGP sessions will be established between PE3, PE4, ASBR2 and ASBR4. eBGP sessions will be established among ASBR1, ASBR2, ASBR3, and ASBR4.

All PEs and ASBRs are enabled for the E-VPN SAFI, and exchange E-VPN Ethernet A-D routes - one route per AC. The ASBRs re-advertise the Ethernet A-D routes with Next Hop attribute set to their IP addresses. The link between the CE and the PE is either a C-tagged or S-tagged interface, as described in [802.1Q], that can carry a single VLAN tag or two nested VLAN tags. This interface is set up as a trunk with multiple VLANs.

A VPWS with multiple sites or multiple EVPL services on the same CE port can be included in one EVI between 2 or more PEs. An Ethernet Tag corresponding to each P2P connection and known to both PEs is used to identify the services multiplexed in the same EVI.

For CE multi-homing, the Ethernet AD Route encodes the ESI associated with the CE. This allows flow-based load-balancing of traffic between PEs connected to the same multi-homed CE. The VPN ID MUST be the same on both PEs attached to the site. The Ethernet Segment route may be used too, for discovery of multi-homed CEs. In all cases traffic follows the transport paths, which may be asymmetric.

4 E-VPN Comparison to PW Signaling

In E-VPN, service endpoint discovery and label signaling are done concurrently using BGP. Whereas, with VPWS based on [RFC4448], label signaling is done via LDP and service endpoint discovery is either through manual provisioning or through BGP. In VPWS, redundancy is limited to Active/Standby mode, while with E-VPN both Active/Active and Active/Standby redundancy modes can be supported. In VPWS, backup PWs are not used to carry traffic, while E-VPN traffic can be load-balanced among primary and secondary PEs. On link or node failure, E-VPN can trigger failover with the withdrawal of a single BGP route per service, whereas with VPWS PW redundancy, the failover sequence requires exchange of two control plane messages: one message to deactivate the group of primary PWs and a second message to activate the group of backup PWs associated with the access link. Finally, E-VPN may employ data plane local repair mechanisms not available in VPWS.

5 ESI Bandwidth Attribute

The ESI Bandwidth Attribute is a new optional BGP attribute that will be associated with the Ethernet AD route used to realize the EVPL services.

```
         +--------------------------------------+
         |       Type     (2 octets)            |
         +--------------------------------------+
         |       Length   (2 octets)            |
         +--------------------------------------+
         |        Flags    (1 Octet)            |
         +--------------------------------------+
         |       Reserved=0(1 Octet)            |
         +--------------------------------------+
         |       Reverse SENDER_TSPEC           |
         +--------------------------------------+
```

The content of the SENDER_TSPEC are as defined in [RFC 2210] section 3.1.

When a PE receives this attribute for a given EVPL it MUST request the appropriate resources described in the SENDER_TSPEC from the PSN towards the other EVPL service destination PE originating the message. When resources are allocated from the PSN for a given EVPL service, then the PSN SHOULD account for the Bandwidth requested by this EVPL service.

In the case where PSN resources are not available, the PE receiving this attribute MUST re-send its local Ethernet AD routes for this EVPL service with the ESI Bandwidth attribute and with the Flags set to 1 "PSN Resources Unavailable".

6 VPWS with multiple sites

The future revision of this draft will describe how a VPWS among multiple sites (full mesh of P2P connections - one per pair of sites) can be setup automatically without any explicit provisioning of P2P connections among the sites.

7 Security Considerations

This document does not introduce any additional security constraints.

8 IANA Considerations

TBD

9 References

9.1 Normative References

[KEYWORDS] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC 2210] Wroclawski, J. "The Use of RSVP with IETF Integrated Services", RFC 2210, September 1997

9.2 Informative References

[EVPN-REQ] A. Sajassi, R. Aggarwal et. al., "Requirements for Ethernet VPN", draft-ietf-l2vpn-evpn-req-00.txt.

[EVPN] A. Sajassi, R. Aggarwal et. al., "BGP MPLS Based Ethernet VPN", draft-ietf-l2vpn-evpn-00.txt.

Authors' Addresses

Sami Boutros
Cisco
170 West Tasman Drive
San Jose, CA 95134, US
Email: sboutros@cisco.com

Ali Sajassi
Cisco
170 West Tasman Drive
San Jose, CA 95134, US
Email: sajassi@cisco.com

Samer Salam
Cisco
595 Burrard Street, Suite 2123
Vancouver, BC V7X 1J1, Canada
Email: ssalam@cisco.com

John Drake
Juniper Networks
Email: jdrake@juniper.net

Jeff Tantsura
Ericsson
Email: jeff.tantsura@ericsson.com