Network Working Group
Internet Draft
Category: Standard Track

Rafi Ram, Orckit-Corrigent
Daniel Cohn, Orckit-Corrigent
Raymond Key, Huawei
P. Agarwal, Broadcom
Yuqun (Sam) Cao, Ruijie Networks
Josh Rogers, TW Cable

Expires: September 5, 2012

Mar 5, 2012

# Extension to VPLS for E-Tree Using Multiple PWs
# draft-ram-l2vpn-etree-multiple-pw-01

Status of this Memo

Abstract

This document proposes a solution for Metro Ethernet Forum (MEF) Ethernet Tree (E-Tree) support in Virtual Private LAN Service using LDP Signaling (LDP-VPLS) [RFC4762],BGP signaling (BGP-VPLS) [RFC4761] or BGP auto-discovery (BGP-AD) [RFC6074]. The proposed solution is characterized by the use of two PWs between a pair of PEs. This solution is applicable for both VPLS and H-VPLS.

Table of Contents

## 1. Introduction

This document proposes a solution for Metro Ethernet Forum (MEF) Tree (E-Tree) support in Virtual Private LAN Service using LDP Signaling (LDP-VPLS) [RFC4762], BGP Signaling (BGP-VPLS) [RFC4761] or BGP auto-discovery (BGP-AD) [RFC6074].

[Draft ETree VPLS Req] is used as requirement specification.

The proposed solution is characterized by the use of two PWs between a pair of PEs, which requires extension to the current VPLS standards [RFC4762],[RFC4761] and [RFC6074].

This solution is applicable for both VPLS and H-VPLS.

The proposed solution is composed of three main components:

- Current VPLS standards: LDP-VPLS [RFC4762],BGP-VPLS [RFC4761] and BGP-AD [RFC6074]
- Extensions to the above specified in this document
- PE local split horizon mechanism

## 2. Conventions used in this document

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].

In this document, these words will appear with that interpretation only when in ALL CAPS. Lower case uses of these words are not to be interpreted as carrying RFC-2119 significance.

## 3. The Problem

[Draft ETree VPLS Req] identifies the problem when there are two or more PEs with both Root AC and Leaf AC.

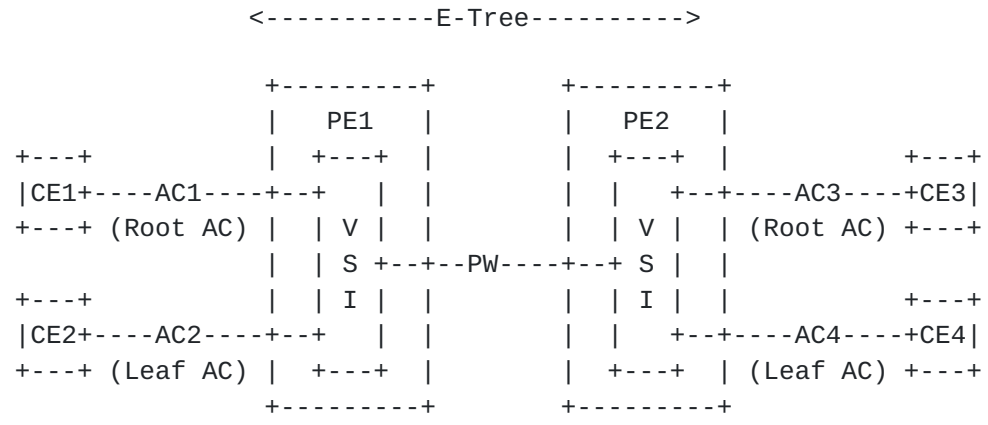



Figure 1: Problem Scenario for Leaf-to-Leaf Communication Restriction

When PE2 receives a frame from PE1 via the Ethernet PW:

- PE2 does not know whether the ingress AC is a Leaf AC or not
- PE2 does not have sufficient information to enforce the Leaf-to-Leaf communication restriction

## 4. The 2-PW Solution

A simple fix is to carry additional information with each frame on the PW, indicating whether the frame is originated from a Leaf AC or a Root AC on the ingress PE.

The proposed solution uses a pair of PWs to interconnect two VPLS PEs:

- First PW is used for frames originated from Root ACs
- Second PW is used for frames originated from Leaf ACs

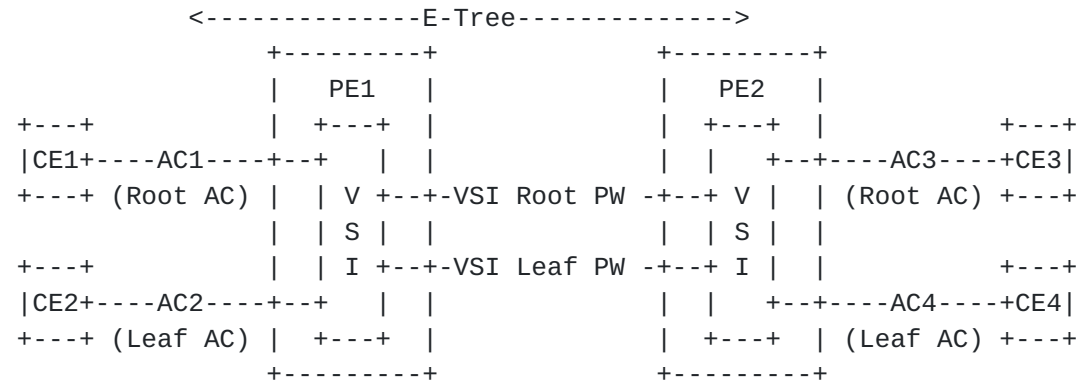



Figure 2: Two-PW Solution for Leaf-to-Leaf Communication Restriction

The next sections specify the required extension to current VPLS standards.

## 5. AC E-Tree Type

Each AC connected to a specific VPLS instance on a PE MUST have an AC E-Tree Type attribute, either Leaf AC or Root AC. For backward compatibility, the default AC E-Tree Type MUST be Root.

This AC E-Tree Type is locally configured on a PE and no signaling is required between PEs.

## 6. Extension to LDP-VPLS for E-Tree

This section specifies extensions to LDP-VPLS [RFC 4762] to support E-Tree requirements. These extensions apply to both FEC types specified in [RFC 4762], namely PWid and generalized PWid.

### 6.1. VSI E-Tree Type and Identifier

Two new PW interface parameters (as defined in section 5.5 of [RFC4447]) are defined for use in E-Tree VPLS: VSI E-Tree type and VSI E-Tree identifier.

VSI E-Tree type can be either root or leaf and identifies VSI root PW and VSI leaf PW respectively, as defined in section 4.

VSI E-tree identifier is a number that is used to identify a pair of root and leaf PW as part of the same logical bridge interface.

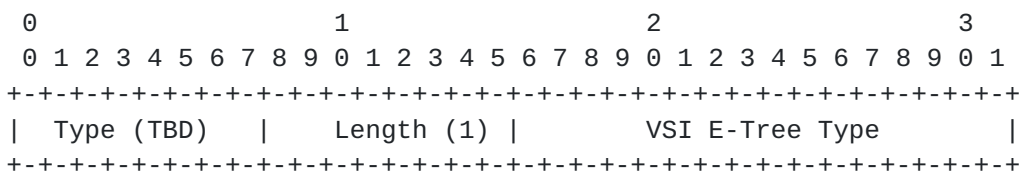
The <VSI E-Tree identifier, VSI E-Tree type> pair SHALL be unique among PWs connecting a pair of VPLS PEs for the same VPLS instance.

#### 6.1.1. VSI E-Tree Type Encoding

The VSI E-Tree type field is encoded as an interface parameters sub-

TLV (as defined in section 5.5 of [RFC4447]).

The field structure is defined as follows:




VSI E-tree Type can take the following values:

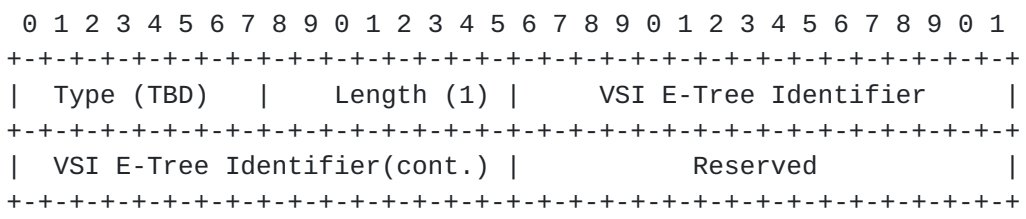
0 E-Tree Root VSI

1 E-Tree Leaf VSI

### 6.1.2. VSI E-Tree Identifier Encoding

The VSI E-Tree identifier field is encoded as an interface parameters sub-TLV (as defined in section 5.5 of [RFC4447]).

The field structure is defined as follows:




VSI E-tree Identifier is a 32-bit number that is used to identify a pair of root and leaf PW as part of the same logical bridge interface, in the context of a pair of VPLS PEs.

The reserved field SHALL be set to zero.

## 6.2. Root/Leaf PWs Signaling

Signaling of root and leaf PWs is required only when two PWs are used for interconnecting between pair of VSIs. As explained in section 6.1:

- Root VSI E-Tree type SHALL be used to signal a root PW.

- Leaf VSI E-Tree type SHALL be used to signal a leaf PW.

PW type signaling rules remain as defined in [RFC4447].

When the generalized PWid encoding (FEC 129) is used, AII shall be set to 1 for the leaf PW.

It should be noted that in a full-mesh VPLS (as opposed to H-VPLS), the following VSI pair types do not require two interconnecting PWs:

Root-only VSI <-> any VSI: only root PW required

Leaf-only VSI <-> leaf-only VSI: no PWs required

Where root-only VSI is a VSI where all ACs are of the root type, and leaf-only VSI is one where all ACs are of the leaf type.

### 6.3. Supporting Remote AC

When PW is used to interconnect between VSI and a remote AC (e.g. the PW1, PW2 in Figure 3), an Ethernet Raw or Ethernet tagged PW types SHALL be used as defined in [RFC4762].

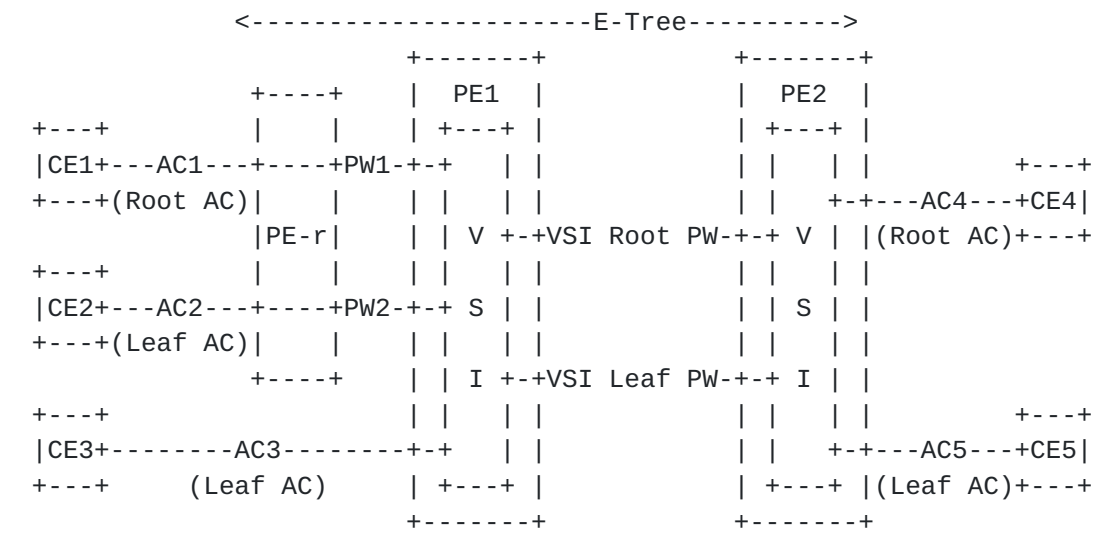



Figure 3: VPLS with Remote AC Connectivity

In addition, the AC type i.e. Root or leaf, SHALL be locally provisioned on the VSI side to specify the remote AC E-Tree Type per PW. Moreover, such PWs that are used for interconnecting between a remote AC and a VSI SHALL considered as separate logical bridge interfaces with respect to MAC address learning/forwarding e.g. traffic forwarding between such PWs is allowed as long as they are not both defined as Leaf.

In Figure 3, AC1 is remotely interconnected to the VPLS service via PW1, and AC2 is remotely interconnected to the VPLS service via PW2.

AC1 is a Root AC and therefore the local type for PW1 in PE1 SHALL

be Root.

AC2 is a Leaf AC and therefore the local type for PW2 in PE1 SHALL be Leaf.

## 7. Extension to BGP-VPLS for E-Tree

This section specifies extensions to BGP-VPLS [RFC 4761] to support E-Tree requirements.

### 7.1. Auto-discovery

Requirements in section 3.2.2 of [RFC 4761] apply, with the following modifications.

A PE SHALL advertise two NLRIs for each E-Tree VPLS instance, with the same VE-ID and non-overlapping label blocks. The PE SHALL indicate that one of the NLRIs signals a root PW and the other one signals a leaf PW by setting the E-Tree type field in the attached Layer2 Info Extended Community, as specified in section 7.2. A special E-tree type is used for the leaf PW when only leaf ACs exist in the VPLS instance, as specified in section 7.2.

### 7.2. PW Setup and Teardown

Requirements in section 3.2.3 of [RFC4761] apply, with the following modifications.

If a PE receives two VPLS NLRI announcements for an E-Tree VPLS instance from a remote PE with the same VE-ID and different root/leaf indication, the PE SHALL set up two PWs to the remote PE.

If a PE with an E-Tree VPLS instance with only leaf ACs receives a VPLS NLRI announcement for this instance from a remote PE with the leaf-only indication, no PWs shall be set up to the remote PE. This rule overrides the previous one.

If a PE receives a legacy VPLS NLRI  for an E-Tree VPLS instance from a remote PE, it will withdraw the Leaf or leaf-only VPLS NLRI it previously advertised and set up only a root PW to the remote PE.

PW setup for each of the PWs follows the rules in 3.2.3 of [RFC4761].

A PW established following the receipt of a VPLS NLRI with root indication will be known as root PW.

A PW established following the receipt of a VPLS NLRI with leaf or leaf-only indication will be known as leaf PW.

Two PWs established following the receipt of VPLS NLRIs with the same VE-ID SHALL be associated to the same logical bridge interface.

### 7.3. Root/Leaf PWs Signaling

The Layer2 Info Extended Community attribute is used to indicate root/leaf assignment for the associated VPLS NLRI.

With reference to Figure 4, bits 4-5 in the control flags are defined for E-Tree type (ET) signaling. Bits C, S have been defined in [RFC4761].

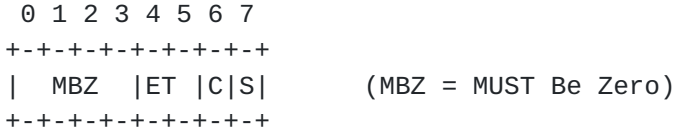



Figure 4 - Control Flags Bit Vector

ET can take the following values:

0 Legacy VPLS NLRI: This PE does not support E-Tree extensions.

1 E-Tree Leaf-only VPLS NLRI: there are only leaf ACs in the VSI, and this is the E-Tree Leaf VPLS NLRI.

2 E-Tree Root VPLS NLRI: there are root ACs in the VSI, and this is the E-Tree Root VPLS NLRI.

3 E-Tree Leaf VPLS NLRI: there are root ACs in the VSI, and this is the E-Tree Leaf VPLS NLRI.

### 7.4. Optimization

As in the LDP case (section 6.2), root and leaf PWs need not be established between every VSI pair. Procedures in this draft avoid the establishment of PWs between leaf-only VSIs, but they do not avoid establishment of leaf PW between root-only VSI and any other VSI. This is a consideration for future versions of the draft.

## 8. Extension to BGP-AD for E-Tree

This section specifies extensions to BGP-AD [RFC6074] to support E-Tree requirements.

### 8.1. Auto-discovery

Requirements in section 3.3.2.1 of [RFC6074] apply, with the following modifications.

Each PE with SHALL advertise two NLRIs for each VPLS instance, with the same VE-ID.

The PE SHALL indicate that one of the NLRIs advertises a root attachment and the other one a leaf attachment by setting the PE_addr field to zero and one respectively in the NLRIs.

### 8.2. PW Setup and Teardown

Requirements in section 3.3.3 of [RFC6074] apply, with the following modifications.

If a PE receives two VPLS NLRI announcements from a remote PE with the same VE-ID and different root/leaf indication, the PE SHALL set up two PWs to the remote PE. PW setup for each of the PWs follows the rules in 3.3.3 of [RFC6074].

A PW established following the receipt of a VPLS NLRI with root indication will be known as root PW.

A PW established following the receipt of a VPLS NLRI with leaf indication will be known as leaf PW.

Two PWs established following the receipt of VPLS NLRIs with the same VE-ID SHALL be associated to the same logical bridge interface.

### 8.3. Optimization

As in the LDP case (section 6.2), root and leaf PWs need not be established between every VSI pair. However, BGP-AD optimization to avoid root or leaf PW setup in these cases is not considered in this draft and is left as a consideration for future versions.

## 9. Data Forwarding Requirements

On frame reception, two PWs associated to the same logical bridge interface SHALL be handled as a single bridge interface with respect to MAC address learning/forwarding, e.g. traffic SHALL NOT be forwarded between such PWs and MAC addresses in frames arriving at any of the PWs SHALL be learned on a common logical bridge interface.

On transmission, the VPLS processing entity SHALL send root-originated traffic via the root PW, and SHALL send leaf-originated traffic via the leaf PW.

An egress PE SHALL NOT deliver a frame originated at a leaf AC to another leaf AC.

The following specifies how AC E-Tree type per frame is determined:

- o A frame received from a root PW indicates that the frame was originated from a root AC.
- o A frame received from a leaf PW indicates that the frame was originated from a leaf AC.
- o For the case where both ingress AC and egress AC are on the same PE, local split horizon implementation on the PE will be sufficient, and is not further discussed in this document.

## 10. Backward Compatibility

### 10.1. LDP-VPLS

Root or leaf VSI E-Tree type and identifier parameters SHALL be used only in cases where both PEs are VPLS capable and both support E-Tree extensions defined in this document.

### 10.2. BGP-VPLS

VPLS NLRIs with root/leaf indication are transmitted only to remote PEs that support E-Tree extensions defined in this document.

### 10.3. BGP-AD

VPLS NLRIs with root/leaf indication SHALL be transmitted only to remote PEs that support E-Tree extensions defined in this document.

## 11. Compliance with Requirements

This refers to [Draft ETree VPLS Req] Section 5 Requirements.

The solution prohibits communication between any two Leaf ACs in a VPLS instance.

The solution allows multiple Root ACs in a VPLS instance.

The solution allows Root AC and Leaf AC of a VPLS instance to co-exist on any PE.

The solution is applicable to LDP-VPLS [RFC4762], BGP-VPLS [RFC4762] and BGP-AD [RFC6074].

The solution is applicable to Case 1: Single technology "VPLS Only".

## 12. Security Considerations

This will be added in later version.

## 13. IANA Considerations

Additional assignments will be required for the new interface parameter sub-TLV types introduced in Section 4.2. Details will be added in a later version.

## 14. Acknowledgements

The authors wish to acknowledge the contributions of Luca Martini and Amir Halperin.

## 15. References

### 15.1. Normative References

[RFC2119] Bradner, S., Key words for use in RFCs to Indicate Requirement Levels, BCP 14, RFC 2119, March 1997.

[RFC4447] Martini, L., and al, Pseudowire Setup and Maintenance Using the Label Distribution Protocol (LDP), April 2006

[RFC4762] Lasserre & Kompella, Virtual Private LAN Service (VPLS) Using Label Distribution Protocol (LDP) Signaling, January 2007

[RFC4761] Rekhter & Kompella, Virtual Private LAN Service (VPLS) Using BGP for Auto-Discovery and Signaling, January 2007

[RFC6074] Rosen, Davie, Radoaca and Luo, Provisioning, Auto-Discovery, and Signaling in Layer 2 Virtual Private Networks (L2VPNs), January 2011

### 15.2. Informative References

[Draft VPLS ETree Req] Key, et al., Requirements for MEF E-Tree Support in VPLS, draft-key-l2vpn-vpls-etree-reqt-04, September 2011

Authors' Addresses

Rafi Ram
Orckit-Corrigent
126 Yigal Alon St.
Tel Aviv, Israel
Email: rafir@orckit.com

Daniel Cohn
Orckit-Corrigent
126 Yigal Alon St.
Tel Aviv, Israel
Email: danielc@orckit.com

Raymond Key
Huawei
Email: raymond.key@ieee.org

Puneet Agarwal
Broadcom
3151 Zanker Road
San Jose, CA 95134
Email: pagarwal@broadcom.com

Yuqun (Sam) Cao
Ruijie Networks
618 Jinshan Road, Fuzhou 350002, China
Email: yuqun.cao@gmail.com

Josh Rogers
Time Warner Cable
11921 N MoPac Expwy
Austin, TX 78759
USA
Email: josh.rogers@twcable.com