Network Working Group
Internet Draft
Category: Standard Track
Expires: October 2012

Simon Delord, Alcatel-Lucent
Raymond Key, Huawei
Frederic Jounay, Orange CH
Wim Henderickx, Alcatel-Lucent
Lucy Yong, Huawei
Lizhong Jin, ZTE

April 16, 2012

# Control Word Reserved bit for use in E-Tree
# draft-delord-pwe3-cw-bit-etree-07

## Status of this Memo



## Abstract

The extension described in this document allows a pair of PEs connected via an Ethernet Pseudowire (PW) to signal via the use of the Control Word (CW) whether the Ethernet frame encapsulated is coming from a Root AC or a Leaf AC. This allows a PE receiving this frame to decide whether it can be forwarded to a Leaf AC or not. Such forward or drop decision is an additional filtering action after the MAC-based forwarding decision. This applies to both P2P PW and P2MP PW.

Table of Contents



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 [RFC2119].

## 1. Introduction

This document proposes to use a specific bit within the Control Word (CW) present on top of an Ethernet PW for a PE to indicate to a remote PE connected via a Pseudowire (PW) whether the Ethernet frame carried comes from a Root AC or a Leaf AC in an E-Tree construct using VPLS. This applies to both P2P PW and P2MP PW.

[RFC4385] describes the preferred design of a Pseudowire Emulation Edge-to-Edge (PWE3) Control Word to be used over an MPLS packet switched network.

[RFC4447] specifies extensions to Label Distribution Protocol (LDP) for PEs to setup PWs. It also describes the procedures for PEs to signal to each other whether the CW is to be used or not.

[RFC4448] specifies the encapsulation of Ethernet/802.3 PDUs within a PW.

[Draft VPLS ETree Req] provides the functional requirements for MEF E-Tree support in VPLS.

## 2. Terminology

E-Tree

An Ethernet VPN in which each Root AC can communicate with every other AC, whereas Leaf ACs can only communicate with Root ACs. Each AC on an E-Tree construct is designated as either a Root AC or a Leaf AC. There can be multiple Root ACs and Leaf ACs per E-Tree construct.

Root AC

An ingress frame at a Root AC can be delivered to one or more of any of the other ACs in the E-Tree. Please note that this AC is bidirectional.

Leaf AC

Ingress frame at a Leaf AC can only be delivered to one or more Root ACs in the E-Tree. Ingress frame at a Leaf AC MUST NOT be delivered to any Leaf ACs in the E-Tree. Please note that this AC is bidirectional.

## 3. Scope and Applicability

### 3.1. Scope & Problem

One important application for carriers is the deployment of E-Tree services over an MPLS backbone using VPLS.

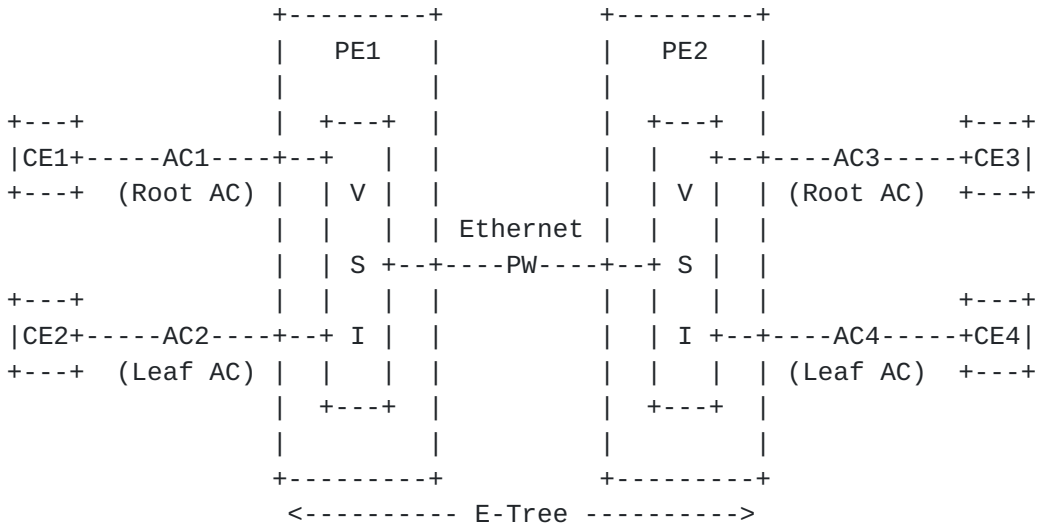



Figure 1 describes a scenario where PE1 and PE2 need to establish an E-Tree construct between different Ethernet endpoints. Each PE has 2 ACs connected to a VSI. These VSIs are then linked together via an Ethernet PW.

AC1 is a Root AC on PE1 and AC2 is a Leaf AC on PE1.

AC3 is a Root AC on PE2 and AC4 is a Leaf AC on PE2.

With an E-Tree construct:

- AC1 can receive from and transmit to AC2, AC3 and AC4
- AC3 can receive from and transmit to AC1, AC2 and AC4
- AC2 and AC4 can receive from and transmit to AC1 and AC3
- AC2 and AC4 cannot receive from or transmit to each other

When an Ethernet Frame is received on PE1 via AC1, the frame can be transmitted to AC2 and via the Ethernet PW connecting PE1 and PE2 to AC3 and AC4.

However when an Ethernet Frame is received on PE1 via AC2, the frame can be transmitted to AC3 via the Ethernet PW but not to AC4.

When PE2 receives an Ethernet frame from PE1 via the Ethernet PW, there is no indication whether it is from a Leaf AC or a Root AC, PE2 is therefore unable to decide whether it can be forwarded to AC4 (Leaf AC) as per the E-Tree construct.

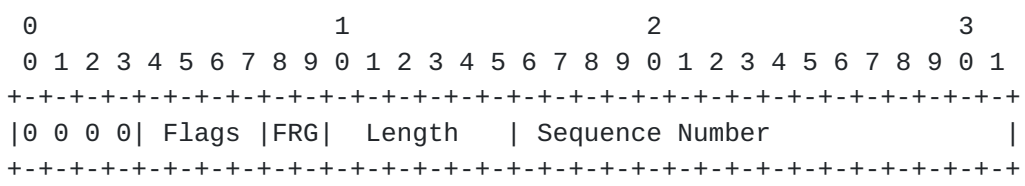
### 3.2. Proposed Approach

A simple fix to this problem is to carry one additional bit of information (0 or 1) for each Ethernet payload frame on the Ethernet PW.

This per-payload signaling approach indicates whether the payload frame is from a Leaf AC or a Root AC on the ingress PE.

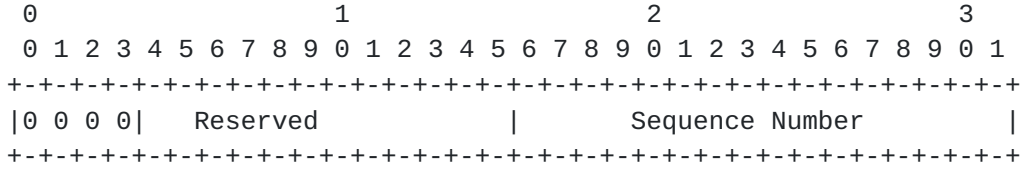
Based on this per-payload extra-information, the egress PE can decide whether the payload frame can be forwarded to a Leaf AC or not.

[RFC4385] defines in section 3 the Preferred PW MPLS Control Word.




Flags (bits 4 to 7):
These bits MAY be used by for per-payload signaling. Their semantics MUST be defined in the PW specification.

[RFC4448] defines in section 4.6 the Control Word for Ethernet PW.




The Flags (bits 4 to 7 of the Control Word) are "reserved for future use" bits in the current Ethernet PW standard that can be used for such per-payload signaling.

### 3.3. Applicability

The proposed approach in this document will allow the egress PE to make a decision for each Ethernet frame arriving on the PW from the ingress PE whether this frame can be forwarded towards a Leaf AC.

Obviously, each PE is also responsible for locally identifying which ACs are Root ACs or Leaf ACs and for not forwarding Ethernet frames between Leaf ACs. The most common procedure for doing so is to apply a split horizon rule between Leaf ACs. However this is out of scope of this document.

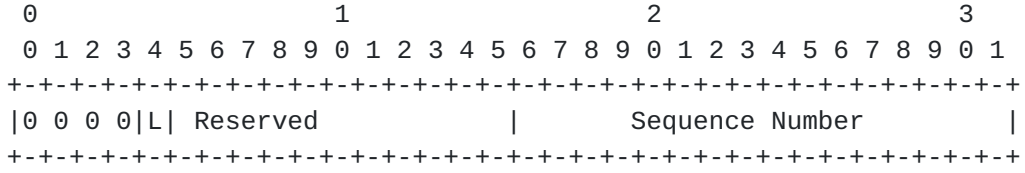
## 4. Control Word Extensions

### 4.1. Ethernet Control Word Reserved Bit

[RFC4448] defines in section 4.6 the CW format for an Ethernet PW.

The Control Word defined in this section is based on the Generic PW MPLS Control Word as defined in [RFC4385]. It provides the ability to sequence individual frames on the PW, avoidance of equal-cost multiple-path load-balancing (ECMP), and Operations and Management (OAM) mechanisms including VCCV.

The Control Word is defined as follows:




In the above diagram, the first 4 bits MUST be set to 0 to indicate PW data.

Leaf bit (L).

The L-bit is used to flag the presence of an Ethernet frame from a Leaf AC as follows:

L = 1 The Ethernet Frame comes from a Leaf AC
L = 0 The Ethernet Frame comes from a Root AC

The rest of the first 16 bits are reserved for future use. They MUST be set to 0 when transmitting, and MUST be ignored upon receipt.

The next 16 bits provide a sequence number that can be used to guarantee ordered frame delivery. The processing of the sequence number field is OPTIONAL.

The sequence number space is a 16-bit, unsigned circular space. The sequence number value 0 is used to indicate that the sequence number check algorithm is not used. The sequence number processing algorithm is found in [RFC4385].

## 5. Procedures

[RFC4447] defines the procedures by which PEs maintain and exchange PW information via LDP.

This document does not change any of the procedures defined in [RFC4447]. The rules for negotiating the use of the CW are defined in section 6.

### 5.1. Detailed PE PW Setup Procedures

In order for a PEs to signal if an Ethernet frame travelling on an Ethernet PW comes from a Root AC or a Leaf AC, it MUST use the extension defined in this document.

The procedure MUST follow the standard procedures described in [RFC4447].

The Label Mapping messages that are sent in order to set up these PWs MUST have c=1. When a Label Mapping message for a PW of one of these types is received and c=0, a Label Release message MUST be sent, with an "Illegal C-bit" status code. In this case, the PW will not be enabled.

### 5.2. Detailed PE Forwarding Procedures

#### 5.2.1. Forwarding PE

The L-bit MUST only be used when the first 4 bits of the CW are equal to 0000.

The L-bit MUST be set to 0 when the first 4 bits of the CW are equal to 0001.

If the Ethernet frame that has to be sent across an Ethernet PW (or set of Ethernet PWs) comes from a Root AC, the L bit MUST be set to 0.

If the Ethernet frame that has to be sent across an Ethernet PW (or set of Ethernet PWs) comes from a Leaf AC, the L bit MUST be set to 1.

#### 5.2.2 Receiving PE

Upon reception of an Ethernet Frame where the L bit in the CW is set to 0, this frame can be forwarded to any AC belonging to this VPN instance.

Upon reception of an Ethernet Frame where the L bit in the CW is set to 1, this frame can be forwarded to any Root AC belonging to this VPN instance but it MUST NOT be forwarded to any Leaf AC. This forward or drop decision is an additional filtering action after the MAC-based forwarding decision.

## 6. Security Considerations

This will be added in later version of this document.

## 7. IANA Considerations

There are no specific IANA considerations in this document.

## 8. Acknowledgments

The authors would like to thank Yuji Kamite for his valuable comments.

## 9. References

### 9.1. Normative References

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

[RFC4385] Bryant,S., Swallow, G., and Al, Pseudowire Emulation Edge-to-Edge (PWE3) Control Word for Use over an MPLS PSN, February 2006.

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

[RFC4448] Martini, L., and al, Encapsulation Methods for Transport of Ethernet over MPLS Networks, April 2006

### 9.2. Informative References

[Draft VPLS ETree Req]
Key, et al., Requirements for MEF E-Tree Support in VPLS, draft-ietf-l2vpn-etree-reqt-00 (work in progress), October 2011

## Authors' Addresses

Simon Delord
Alcatel-Lucent
Email: simon.delord@gmail.com

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

Frederic Jounay
Orange CH
4 rue caudray 1020 Renens
Switzerland
Email: frederic.jounay@orange.ch

Wim Henderickx
Alcatel-Lucent
Copernicuslaan 50
2018 Antwerp, Belgium
Email: wim.henderickx@alcatel-lucent.com

Lucy Yong
Huawei USA
1700 Alma Dr. Suite 500
Plano, TX 75075, USA
Email: lucy.yong@huawei.com

Lizhong Jin
ZTE Corporation
889, Bibo Road
Shanghai, 201203, China
Email: lizhong.jin@zte.com.cn

## Copyright Notice