Encapsulation Methods for Transport of PPP/HDLC Over IP and MPLS Networks

draft-ietf-pwe3-hdlc-ppp-encap-mpls-01.txt

Status of this Memo

This document is an Internet-Draft and is in full conformance with all provisions of Section 10 of RFC2026.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

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."

The list of current Internet-Drafts can be accessed at http://www.ietf.org/ietf/1id-abstracts.txt.

The list of Internet-Draft Shadow Directories can be accessed at http://www.ietf.org/shadow.html.

Abstract

A Pseudowire (PW) can be used to carry PPP, or HDLC Protocol Data Units over an IP or MPLS network without terminating the PPP/HDLC protocol. This enables service providers to offer "emulated" HDLC, or PPP link services over existing IP or MPLS networks. This document specifies the encapsulation of PPP/HDLC PDUs within a pseudowire.
1. Specification of Requirements

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.

2. Introduction

A PPP/HDLC Pseudowire (PW) allows PPP/HDLC Protocol Data Units (PDUs) to be carried over an IP network or an MPLS network. In addressing the issues associated with carrying a PPP/HDLC PDU over a PSN, this document assumes that a Pseudowire (PW) has been set up by some means outside the scope of this document. This may be via manual configuration, or a signaling protocol such as that defined in [1] or [7]. As described in [8], this PW may be tunneled through an MPLS, IPv4 or IPv6 PSN.

The following figure describe the reference models which are derived from [8] to support the HDLC/PPP PW emulated services.
This document specifies the emulated PW encapsulation for PPP, and HDLC. Although different layer 2 protocols require different information to be carried in this encapsulation, an attempt has been made to make the encapsulation as common as possible for all layer 2 protocols. Other layer 2 protocols are described in separate documents. [4] [5] [6]

This document also specifies the way in which the demultiplexer field is added to the emulated PW encapsulation when an MPLS label is used as the demultiplexer field. QoS related issues are not discussed in this document. For the purpose of this document PE1 will be defined as the ingress router, and PE2 as the egress router. A layer 2 PDU will be received at PE1, encapsulated at PE1, transported, decapsulated at PE2, and transmitted out of PE2.
3. General encapsulation method

3.1. The Control Word

There are three requirements that may need to be satisfied when transporting layer 2 protocols over an IP or MPLS backbone:

- i. Sequentiality may need to be preserved.
- ii. Small packets may need to be padded in order to be transmitted on a medium where the minimum transport unit is larger than the actual packet size.
- iii. Control bits carried in the header of the layer 2 frame may need to be transported.

When carrying HDLC/PPP over an IP or MPLS backbone sequentiality may need to be preserved. The OPTIONAL control word defined here addresses this requirement. Implementations MUST support sending no control word, and MAY support sending a control word.

In all cases the egress router must be aware of whether the ingress router will send a control word over a specific virtual circuit. This may be achieved by configuration of the routers, or by signaling, for example as defined in [1]. The control word is defined as follows:

```
+-------------------------------+-------------------------------+-------------------------------+-------------------------------+
|         |         |         |         |
| 0 0 0 0 | 0 0 0 0 | B E |   Length |     Sequence Number           |
+-------------------------------+-------------------------------+-------------------------------+-------------------------------+
```

Figure 2: MPLS PWE3 Control Word

In the above diagram the first 4 bits are the PID as defined in [8]. B and E are fragmentation bits and their functionality is specified in [9].

The next 4 bits provide space for carrying protocol specific flags. These are not used for HDLC/PPP and they MUST be set to 0 when transmitting, and MUST be ignored upon receipt.

The next 6 bits provide a length field, which is used as follows: If the packet’s length (defined as the length of the layer 2 payload plus the length of the control word) is less than 64 bytes, the length field MUST be set to the packet’s length. Otherwise the length field MUST be set to zero. The value of the length field, if non-zero, can be used to remove any padding. When the packet reaches the
service provider’s egress router, it may be desirable to remove the padding before forwarding the packet.

The next 16 bits provide a sequence number that can be used to guarantee ordered packet 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 an unsequenced packet.

3.1.1. Setting the sequence number

For a given PW, and a pair of routers PE1 and PE2, if PE1 supports frame sequencing then the following procedures should be used:

- the initial frame transmitted on the PW MUST use sequence number 1
- subsequent frames MUST increment the sequence number by one for each frame
- when the transmit sequence number reaches the maximum 16 bit value (65535) the sequence number MUST wrap to 1

If the transmitting router PE1 does not support sequence number processing, then the sequence number field in the control word MUST be set to 0.

3.1.2. Processing the sequence number

If a router PE2 supports receive sequence number processing, then the following procedures should be used:

When a PW is initially set up, the "expected sequence number" associated with it MUST be initialized to 1.

When a frame is received on that PW, the sequence number should be processed as follows:

- if the sequence number on the frame is 0, then the frame passes the sequence number check
- otherwise if the frame sequence number >= the expected sequence number and the frame sequence number - the expected sequence number < 32768, then the frame is in order.
- otherwise if the frame sequence number < the expected sequence number and the expected sequence number - the frame sequence number >= 32768, then the frame is in order.
- otherwise the frame is out of order.

If a packet is in order then, it can be delivered immediately. If the packet is in order, then the expected sequence number MUST be set using the algorithm:

```plaintext
expected_sequence_number := frame_sequence_number + 1 mod 2**16
if (expected_sequence_number = 0) then expected_sequence_number := 1;
```

Packets which are received out of order MAY be dropped or reordered at the discretion of the receiver.

A simple extension of the above processing algorithm can be used to detect lost packets.

If a router PE2 does not support receive sequence number processing, then the sequence number field MAY be ignored.

3.2. MTU Requirements

The network MUST be configured with an MTU that is sufficient to transport the largest encapsulation frames. If MPLS is used as the tunneling protocol, for example, this is likely to be 12 or more bytes greater than the largest frame size. Other tunneling protocols may have longer headers and require larger MTUs. If the ingress router determines that an encapsulated layer 2 PDU exceeds the MTU of the tunnel through which it must be sent, the PDU MUST be dropped. If an egress router receives an encapsulated layer 2 PDU whose payload length (i.e., the length of the PDU itself without any of the encapsulation headers), exceeds the MTU of the destination layer 2 interface, the PDU MUST be dropped.

4. Protocol-Specific Details
4.1. HDLC

HDLC mode provides port to port transport of HDLC encapsulated traffic. The HDLC PDU is transported in its entirety, including the HDLC address, control and protocol fields, but excluding HDLC flags and the FCS. Bit/Byte stuffing is undone. The control word is OPTIONAL. If the control word is used then the flag bits in the control word are not used, and MUST be set to 0 when transmitting, and MUST be ignored upon receipt.

The HDLC mode is suitable for port to port transport of Frame Relay UNI or NNI traffic. It must be noted, however, that this mode is transparent to the FECN, BECN and DE bits.

4.2. PPP

PPP mode provides point to point transport of PPP encapsulated traffic, as specified in [3]. The PPP PDU is transported in its entirety, including the protocol field (whether compressed using PFC or not), but excluding any media-specific framing information, such as HDLC address and control fields or FCS. Since media-specific framing is not carried the following options will not operate correctly if the PPP peers attempt to negotiate them:

- Frame Check Sequence (FCS) Alternatives
- Address-and-Control-Field-Compression (ACFC)
- Asynchronous-Control-Character-Map (ACCM)

Note also that PW LSP Interface MTU negotiation as specified in [1] is not affected by PPP MRU advertisement. Thus if a PPP peer sends a PDU with a length in excess of that negotiated for the PW tunnel that PDU will be discarded by the ingress router.

The control word is OPTIONAL. If the control word is used then the flag bits in the control word are not used, and MUST be set to 0 when transmitting, and MUST be ignored upon receipt.

5. Using an MPLS Label as the Demultiplexer Field

To use an MPLS label as the demultiplexer field, a 32-bit label stack entry [2] is simply prepended to the emulated PW encapsulation, and hence will appear as the bottom label of an MPLS label stack. This label may be called the "PW label". The particular emulated pseudo-wire identified by a particular label value must be agreed by the ingress and egress LSRs, either by signaling (e.g., via the methods of [1]) or by configuration. Other fields of the label stack entry are
set as follows.

5.1. MPLS Shim EXP Bit Values

If it is desired to carry Quality of Service information, the Quality of Service information SHOULD be represented in the EXP field of the PW label. If more than one MPLS label is imposed by the ingress LSR, the EXP field of any labels higher in the stack SHOULD also carry the same value.

5.2. MPLS Shim S Bit Value

The ingress LSR, PE1, MUST set the S bit of the PW label to a value of 1 to denote that the PW label is at the bottom of the stack.

6. Security Considerations

This document specifies only encapsulations, and not the protocols used to carry the encapsulated packets across the network. Each such protocol may have its own set of security issues, but those issues are not affected by the encapsulations specified herein.

7. Intellectual Property Disclaimer

This document is being submitted for use in IETF standards discussions.

8. References


IP/MPLS Networks", draft-ietf-pwe3-ethernet-encap-01.txt. (work in progress)


9. Author Information

Luca Martini
Cisco Systems, Inc.
9155 East Nichols Avenue, Suite 400
Englewood, CO, 80112
e-mail: lmartini@cisco.com

Giles Heron
Tellabs
Abbey Place
24-28 Easton Street
High Wycombe
Bucks
HP11 1NT
UK
e-mail: giles.heron@tellabs.com
Eric C. Rosen
Cisco Systems, Inc.
1414 Massachusetts Avenue
Boxborough, MA 01719
E-mail: erosen@cisco.com

Toby Smith
Laurel Networks, Inc.
Omega Corporate Center
1300 Omega Drive
Pittsburgh, PA 15205
e-mail: tob@laurelnetworks.com

Andrew G. Malis
Tellabs
90 Rio Robles Dr.
San Jose, CA 95134
e-mail: Andy.Malis@tellabs.com

Yeongil Seo
463-1 KT Technology Lab
Jeonmin-dong Yusung-gu
Daejeon, Korea
e-mail: syi1@kt.co.kr