

Internet Draft  
Document: <[draft-bryant-pwe3-fr-encap-01.txt](#)>  
Expires: September 2002

Stewart Bryant  
Lloyd Wood  
George Wilkie  
Cisco Systems

March 2002

## A Proposed Frame Relay Encapsulation for PWE3

### 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/lid-abstracts.txt> The list of Internet-Draft Shadow Directories can be accessed at <http://www.ietf.org/shadow.html>.

### Abstract

This draft describes a pseudo-wire emulation edge-to-edge (PWE3) frame relay encapsulation that conforms to the PWE3 Protocol Layering proposed by Bryant et al [[LAYER](#)], and discusses the services required from the underlaying PWE3 layers.

## Table of Contents

<a href="#">1.</a>	Introduction.....	<a href="#">2</a>
<a href="#">2.</a>	Terminology.....	<a href="#">2</a>
<a href="#">3.</a>	PW Encapsulation.....	<a href="#">3</a>
<a href="#">3.1</a>	Payload Convergence Sub-layer.....	<a href="#">3</a>
<a href="#">3.2</a>	Payload Independent PW Encapsulation Layers.....	<a href="#">4</a>
<a href="#">4.</a>	PSN Tunnel Layer Requirements.....	<a href="#">5</a>
<a href="#">4.1</a>	Multiplexing.....	<a href="#">5</a>
<a href="#">4.2</a>	Segmentation and Reassembly.....	<a href="#">5</a>
<a href="#">4.3</a>	Length and Delivery.....	<a href="#">5</a>
<a href="#">5.</a>	Control Plane.....	<a href="#">6</a>
<a href="#">6.</a>	IANA considerations.....	<a href="#">6</a>
<a href="#">7.</a>	Security.....	<a href="#">6</a>
<a href="#">Appendix A:</a>	Intermediate Formats Considered Harmful.....	<a href="#">7</a>

[1.](#) Introduction

A frame relay pseudo-wire allows frame relay protocol control information and payloads to be carried over a Packet Switched Network. This draft describes how our preferred frame relay encapsulation is carried over the proposed PWE3 protocol layering model [[LAYER](#)], and discusses the services required from the underlying PWE3 layers.

[2.](#) Terminology

The following definitions are used in t his document:

BE CN	Backward Explicit Congestion Notification bit in frame relay header.
-------	---

INTERNET DRAFT

[draft-bryant-pwe3-fr-encap-01](#)

March 2002

Customer Edge (CE)	A device where one end of a service originates and terminates. The CE is not aware that it is using an emulated service rather than a native service.
DE	Discard Eligibility bit in frame relay header.
Native Service Processing (NSP)	Processing of the data received by the PE from the CE before presentation to the PW for transmission across the core.
FECN	Forward Explicit Congestion Notification bit in frame relay header.
Packet Switched Network (PSN)	A network using IP or MPLS as the mechanism of packet forwarding.
Protocol Data Unit (PDU)	The unit of data output to, or received from, the network by a protocol layer.
Provider Edge (PE)	A device that provides PWE3 to a CE.
Pseudo Wire (PW)	A connection between two PEs carried over PSN.
Pseudo Wire Emulation Edge to Edge (PWE3)	A mechanism that emulates the essential attributes of service (such as a T1 leased line or frame relay) over a PSN.
PSN Tunnel	A tunnel across a PSN inside which one or more PWs can be carried.
PWE3 Payload Independent Sub-layers	The protocol sub-layers within PWE3 responsible for sequencing and timing.

### [3.](#) PW Encapsulation

#### [3.1](#) Payload Convergence Sub-layer

Frame Relay uses the PWE3 generic packet payload service [[LAYER](#)], employing the principle of minimum intervention. The native frame relay PDU is transported as received, excluding only HDLC flags and the frame-check sequence (FCS). Bit stuffing is undone. (This

contrasts with the proposal of Kawa et al [[KAWA](#)], which uses an intermediate PW format.)

If the underlying PSN provides a congestion notification service, the congestion state must to be passed up to the Payload Convergence Sub-layer by the intermediate layers. The Payload Convergence Sub-layer may set the frame relay Forward Explicit Congestion Notification (FECN) bit if the packet has experienced congestion traveling across the PSN to the destination PE. It may also set the Backward Explicit Congestion Notification (BECN) bit on frame relay packets on the reverse path of the same PW.

The encapsulation layer may request the use of different PSN QoS settings, depending on the setting of the Discard Eligibility (DE) bit in the frame relay header.

The Emulated Service (as defined in figure 1 of [[LAYER](#)]) is responsible for providing the following frame relay service policies.

- o CIR (Committed Information Rate) or throughput.
- o Bc (Committed burst size).
- o Be (Excess Burst Size).
- o Maximum frame size.

Typically, the NSP would be responsible for policing the traffic to CIR, Bc, Be before passing to the PW. Some indication of the committed burst size (which should get guaranteed delivery) and the excess burst size (which may be delivered with lower probability) would be provided by the NSP to the PW. The PW and the remote NSP are then responsible for ensuring that committed traffic arrives at the remote CE. In other implementations, typically those operating over a high bandwidth core, all the service policy processing could be

delegated to the destination PE.

If the frame relay PDU conforms to the maximum frame size, but the PW encapsulated frame relay PDU exceeds the PSN maximum PDU length, the segmentation and reassembly considerations described in [[LAYER](#)] are followed.

### [3.2](#) Payload Independent PW Encapsulation Layers

Frame Relay requires the use of sequencing from the payload independent encapsulation layer to maintain packet order, which is an invariant of Frame Relay networks.

In some cases, the protocol carried over the frame relay VC is known to be insensitive to packet order (for example if the traffic is KNOWN to be IP), in which case there may be a desire to reduce the

overhead associated with the transmission and receive processing of ordered packets. The use of the sequencing service is therefore optional.

The encapsulation of frame relay has no timed delivery or clock recovery requirements. It therefore does not use the PWE3 Timing Sublayer.

## [4.](#) PSN Tunnel Layer Requirements

### [4.1](#) Multiplexing

The frame relay Encapsulation Layer depends on the multiplexing functionality in the PSN Tunnel Layer to provide multiple PWs between PEs. There are two types of frame relay PW:

- o Trunks, in which multiple frame relay VCs are sent over the PW. The normal usage of this is where all traffic on a physical interface to the PE is sent over the PW to a corresponding physical interface on the destination PE.

A PE may contain an NSP that provides a frame relay switch

function, and presents a frame relay trunk to the PW Encapsulation Layer via a virtual interface.

- o Single DLCI, in which the PW carries just the traffic of a single VC.

#### [4.2](#) Segmentation and Reassembly

It is desirable that the MTU of the frame-relay packets received from the CE is constrained so that the packets can be carried over the PSN without fragmentation. If this is not possible, the segmentation and reassembly service of the underlying PSN is used.

#### [4.3](#) Length and Delivery

The underlying PSN is responsible for determining the correct length of a frame relay PDU and delivering the PDU to the PSN Tunnel Layer.

### [5.](#) Control Plane

Mechanisms are needed to set up and tear down frame relay PWs and to carry PVC status across the PW.

An example of the necessary extensions needed to support a frame relay PW across an L2TP PSN tunnel is given in [[TOWNSLEY](#)].

Aspects of the PW control plane specific to frame relay form a work area that needs further study.

### [6.](#) IANA considerations

IANA considerations are for further study.

## 7. Security

PWE3 provides no means of protecting the contents or delivery of the PDU's on behalf of the native service. PWE3 may, however, leverage security mechanisms provided by the PSN Tunnel Layer.

A more detailed discussion of PW security is give in [[LAYER](#)].

## References

Internet-drafts are works in progress available from  
<<http://www.ietf.org/internet-drafts/>>

[KAWA] Kawa et al, Frame relay over Pseudo-Wire Emulation Edge-to-Edge,  
<[draft-kamapabhava-fr-pwe3-00.txt](#)>, work in progress

[LAYER] Bryant et al, Protocol Layering in PWE3,  
<[draft-bryant-pwe3-protocol-layering-01.txt](#)>, work in progress

[MARTINI] L. Martini, Tappan, D., et al. 'Encapsulation Methods for Transport of Layer 2 Frames Over IP and MPLS Networks', work in progress as an internet-draft:  
<[draft-martini-l2circuit-encap-mpls-03.txt](#)>, work in progress

Bryant, et al.

Informational

[Page 6]

---

INTERNET DRAFT

[draft-bryant-pwe3-fr-encap-01](#)

March 2002

[Q.922] ITU-T Recommendation Q.922, ISDN Data Link Layer Specification for Frame Mode Bearer Services, ITU, Geneva, 1992.

[Q933] ITU-T Recommendation Q.933, ISDN Signaling Specification for Frame Mode Bearer Services, Geneva, October 1995.

[TOWNSLEY] Townsley et al, Frame-Relay Pseudo-Wire Extensions for L2TP <[draft-ietf-l2tpext-pwe3-fr-00.txt](#)>, work in progress





```
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
```

The two-octet frame relay header in normal IETF (DoD) notation is:

```

0                               1
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| hi dlci   |C|0|lo dlci|F|B|D|1|
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

where:

C = FR frame C/R (command/response) bit [[Q.922](#)].

F = FECN (Forward Explicit Congestion Notification) bit [[Q.922](#)].

B = BECN (Backward Explicit Congestion Notification) bit [[Q.922](#)].

D = DE indicates the Discard Eligibility [[Q922](#)].

Frame relay header bits 7 and 15 are the Q.922 EA (Extended Address) bits used for address field extension. In the unextended two-octet header these are defined to be 0 and 1 respectively.

Processors rarely have efficient bit manipulation operations, so you cannot normally just assign the bits to their new positions easily. It therefore becomes necessary to isolate each of the DBFC bits in the control word using an AND operation, to use a shift operator to move each bit to its correct position in the frame relay header, and then to load the isolated bit into the header buffer using an OR operator. This requires four operations on each of four bits:

- Move first octet of control word to temporary location.
- AND to isolate bit.
- Shift bit to corresponding location in frame relay header octet.
- Write or OR bit into frame relay header.

A similar number of operations is needed in constructing the Control Word from the received frame relay header. For comparison we can regard this construction as having a cost of: (4 bits \* 4 operations) + (4 bits \* 4 operations) = 32 operations.

Note that we have not included the processing of the DLCI bits (normally an OR operation) in this analysis, since that is an overhead common to all transmission formats. Also note that, for the

purposes of frame relay header manipulation, the EA bits can normally be included in the pro-forma DLCI definition.

## [A.2](#) The Kawa encapsulation

As with the [\[MARTINI\]](#) approach, the frame relay payload-dependent bits are copied to fields in a control word, the frame relay header is stripped from the frame relay PDU, and the header is reconstructed at the far end of the pseudo-wire. Again, the reason for doing this is not given. It has been proposed that [\[KAWA\]](#) be changed to follow the [\[MARTINI\]](#) control word, but at the time of writing this updated approach has not yet been published.

The grouping and ordering of the payload-dependent bits in the [\[KAWA\]](#) control word is more consistent with Q.922, allowing them to be processed as a three bit group (FBD) and a single bit (C). This reduces the processing overhead in comparison with [\[MARTINI\]](#). However, the [\[KAWA\]](#) draft fails to take advantage of the additional performance gains to be made by closely following the layout specified in [\[Q.922\]](#).

The Kawa control word is:

```

      0                               1                               2                               3
    0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| Res |P|F|B|D|C|X|Y|  Length  | Sequence Number                    |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

Again for comparison, the two-octet frame relay header in normal IETF (DoD) notation is:

```

      0                               1
    0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| hi dlci  |C|0|lo dlci|F|B|D|1|
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

In this case we can isolate the FBD bits as group, and knowing that they are in the same position and order in the octet in both Control Word and the frame relay header, we can write them directly to the header buffer in three operations:

- Move first octet of control word to temporary location.
- AND to isolate FBD bits.
- Write bits into frame relay header.

The C bit requires isolation with an AND operator, shifting and

INTERNET DRAFT

[draft-bryant-pwe3-fr-encap-01](#)

March 2002

writing into the header buffer (4 operations).

- Move second octet of control word to temporary location.
- AND to isolate C bit.
- Shift bit to corresponding location in frame relay header octet.
- Write C bit into frame relay header.

This same number of operations is needed in construction of the control word making a comparative cost of:  $(3 + 4) + (3 + 4)$  operations = 14 operations.

14 operations compares well with the 32 operations needed for the [\[MARTINI\]](#) implementation.

### [A.3](#) Cost of using PWE3 Generic Packet Payload

The use of the general pseudo-wire encapsulation technique as proposed in this document has a frame relay header bit processing cost of zero. It is therefore the most processor-efficient approach.

### [A.4](#) Frame Relay Header Translation

A common justification for the use of an intermediate format for the transmission of a frame relay datagram across a PW is the need to translate the header from one format to another.

There are two frame relay header formats in common use: the two-octet version used in the examples in this draft, and the four-octet version.

If the intermediate format is used, the encapsulator has to implement two cases (two-octet to intermediate and four-octet to intermediate), and the decapsulator has to implement two cases (intermediate to two-octet, and intermediate to four-octet). If the frame is transmitted across the pseudo-wire un-translated, then there are three translation cases (Nothing, 2->4 and 4->2). This results in a net reduction in translation and implementation complexity, and an increase in performance.

It is useful to review the memory organization of the two-octet and

four-octet frame relay headers, to fully appreciate the complexity of the translation operation. In normal IETF notation, the two-octet frame relay header is:

```

      0                               1
    0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5
+---+---+---+---+---+---+---+---+
| hi dlci   |C|0|lo dlci|F|B|D|1|
+---+---+---+---+---+---+---+---+

```

and the four-octet frame relay header is

```

      0                               1                               2                               3
    0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| hi dlci   |C|0| dlci |F|B|D|0|   dlci       |0| dlci_lo  |0|1|
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

Bit 30 is the D/C bit. This is set to zero to indicate that the last octet contains DLCI, rather than control, information.

As can be seen by inspecting these definitions, the translation of a two-octet frame relay header to a four-octet frame relay header is a relatively simple operation requiring.

- Move first word of the header to temporary location.
- AND to clear bit 15 (EA = 0).
- Write word into frame relay header at new buffer offset.

Correct setting of the remaining EA bits (bits 23 and 31) and the DLCI Core indicator (D/C) bit (bit 30) can be considered an integral part of writing the least significant thirteen bits of the DLCI.

Translation from a two-octet frame relay header to a four-octet frame relay header has similar complexity:

- Move first word of the header to temporary location.

- OR to set bit 15 (EA = 0).
- Write word into frame relay header at new buffer offset.

Given the simplicity of the frame relay header translation process, they would seem to be no justification for the use of an intermediate PW format for the frame relay header.

#### Authors' Addresses

Stewart Bryant  
Cisco Systems,  
4, The Square,  
Stockley Park,  
Uxbridge UB11 1BL,  
United Kingdom.

Email: [stbryant@cisco.com](mailto:stbryant@cisco.com)  
Phone: +44-20-8756-8828

George Wilkie  
Cisco Systems  
96 Commercial Street,  
Leith,  
Edinburgh, EH6 6LX  
United Kingdom

Email: [gwilkie@cisco.com](mailto:gwilkie@cisco.com)

Lloyd Wood  
Cisco Systems,  
4, The Square,  
Stockley Park,  
Uxbridge UB11 1BL,  
United Kingdom.

Email: [lwood@cisco.com](mailto:lwood@cisco.com)  
Phone: +44-20-8734-4236

#### Full copyright statement

Copyright (C) The Internet Society (2002). All Rights Reserved.

This document and translations of it may be copied and furnished to others, and derivative works that comment on or otherwise explain it or assist in its implementation may be prepared, copied, published and distributed, in whole or in part, without restriction of any kind, provided that the above copyright notice and this paragraph are included on all such copies and derivative works. However, this document itself may not be modified in any way, such as by removing the copyright notice or references to the Internet Society or other Internet organizations, except as needed for the purpose of developing Internet standards in which case the procedures for copyrights defined in the Internet Standards process must be followed, or as required to translate it into languages other than English.

The limited permissions granted above are perpetual and will not be

revoked by the Internet Society or its successors or assigns.

This document and the information contained herein is provided on an "AS IS" basis and THE INTERNET SOCIETY AND THE INTERNET ENGINEERING TASK FORCE DISCLAIMS ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.