Multi-Service Interworking – Ethernet over MPLS

IP/MPLS Forum 12.0.1

IP/MPLS Forum Technical Committee
February 2009

Note: The user's attention is called to the possibility that implementation of the MPLS implementation agreement contained herein may require the use of inventions covered by patent rights held by third parties. By publication of this MPLS implementation agreement the IP/MPLS Forum makes no representation that the implementation of the specification will not infringe on any third party rights. The IP/MPLS Forum take no position with respect to any claim that has been or may be asserted by any third party, the validity of any patent rights related to any such claims, or the extent to which a license to use any such rights may not be available.

Editor:

Rao Cherukuri
Juniper Networks, Inc.

Ali Sajassi
Cisco Systems, Inc.

For more information contact:

The IP/MPLS Forum
48377 Fremont Blvd.
Suite 117
Fremont, CA 94538 USA

Phone: +1 (510) 492-4055
FAX: +1 (510) 492-4001
E-Mail: info@ipmplsforum.org
WWW: http://www.ipmplsforum.org/

Revision History

Version	Change	Date
MFA 12.0.0	Initial version	
IP/MPLS 12.0.1	Minor changes - Clarification and new text for Ethernet frame format to forward BPDU frames	Feb 2009

Table of Contents

1 INTRODUCTION 3

1.1 PURPOSE 3
1.2 SCOPE 3
1.3 OVERVIEW 3

2 DEFINITIONS AND TERMINOLOGY 4

2.1 DEFINITIONS 4
2.2 ACRONYMS AND ABBREVIATIONS 4
2.3 REFERENCE 7
2.3.1 Normative References 7
2.3.2 Informative References 8

3 REFERENCE MODEL 8

4 INTERWORKING MODELS 10

4.1 TWO-SIDED MODEL 10
4.2 CE CONFIGURATION CHANGES 11

5 ETHERNET SERVICE INSTANCE & CLASS OF SERVICE 12

5.1 ETHERNET SERVICE INSTANCE - ESI 12
5.2 ETHERNET CLASS OF SERVICE - COS 16
5.3 ETHERNET PSEUDO WIRE TYPES 18

6 ETHERNET SERVICE INTERWORKING – ENCAPSULATION FORMAT 20

6.1 FRAME RELAY <-> ETHERNET INTERWORKING 20
6.1.1 Frame Formatting and Delimiting 23
6.2 ATM TO ETHERNET INTERWORKING 24
6.2.1 Frame Formatting and Delimiting 26
6.3 FR TO ATM ETHERNET ENCAPSULATION INTERWORKING 27
6.4 PPP TO ETHERNET BRIDGE INTERWORKING 28

7 TRAFFIC MANAGEMENT INTERWORKING 30

8 PVC MANAGEMENT INTERWORKING 30

8.1 FRAME RELAY - ETHERNET INTERWORKING 30
8.1.1 Attachment Circuit Defect Entry/Exit Procedures 30
8.1.2 Pseudo Wire Entry/Exit Procedures 30
8.2 ATM - ETHERNET INTERWORKING 30
8.2.1 Attachment Circuit Defect Entry/Exit Procedures 30
8.2.2 Pseudo Wire Entry/Exit Procedures 30
8.3 FRAME RELAY - ATM INTERWORKING 31
8.3.1 PE attached to Frame Relay AC 31
8.3.2 PE attached to ATM AC 31

ANNEX A 32

A.1 ATM AND ETHERNET ATTACHMENT CIRCUITS WITH ATM PW ENCAPSULATION 32
A.1.1 Frame Formatting and Delimiting 33
A.1.2 TRAFFIC MANAGEMENT AND QOS 33
A.1.3 PVC MANAGEMENT INTERWORKING 34

APPENDIX I 35

I.1 ETHERNET AND FRAME RELAY TM INTERWORKING 35
I.1.1 Mapping of Service Parameters 35

I.1.2 Mapping of Discard Eligibility Indication 37
I.1.3 Overhead Calculation 37
I.2 ETHERNET AND ATM TM INTERWORKING 38
I.2.1 Mapping of Service Parameters 38
I.2.2 Mapping of Drop Precedence Indication: 40
I.2.3 Overhead Calculation 41

1 Introduction

1.1 Purpose

This Specification addresses multi-service interworking of Ethernet over an MPLS core network. The term “multi-service interworking” means that Ethernet frames are transported across the MPLS core in a pseudo wire (PW) [RFC 3985], and that the two attachment circuits (ACs) associated with the PW employ different Layer 2 technologies (e.g. ATM on one AC and Ethernet on the other).

The native service provided to end-users is Ethernet encapsulation. This interworking specification enables carriers and service providers to introduce new Ethernet services, while preserving existing infrastructure like ATM and Frame Relay.

The goal of Ethernet multi-service interworking is to facilitate Ethernet service between customer sites that use different attachment circuit types. It is applicable to both point-to-point Virtual Private Networks (i.e. Virtual Private Wire Service - VPWS) and multipoint VPNs (i.e. Virtual Private LAN Service - VPLS) as defined in the L2VPN framework document [L2 Frame]. It is also applicable to other Ethernet service types, for example, those defined by the Metro Ethernet Forum in [MEF 6]. Since the end-to-end service is of type Ethernet, the encapsulation over a non-Ethernet AC (e.g., ATM) must be of type Ethernet (also known as bridged encapsulation).

Ethernet multi-service interworking has the following characteristics:

- Support for both point-to-point and multipoint connectivity over heterogeneous attachment circuits
- There is no requirement for the PE devices to participate in the address resolution procedures of the CE devices
- Independence from Layer 3 protocols run on CE devices
- Support of different Layer 3 protocols among CE devices simultaneously

1.2 Scope

This document covers Ethernet multi-service interworking among the following combinations of AC types:

1) ATM or FR or PPP to Ethernet
2) ATM to FR
3) ATM or FR to PPP

For an AC of type ATM or FR, only Permanent Virtual Connections (PVCs) are within the scope of this specification; Switched Virtual Connections (SVCs) are outside the scope. Furthermore, this IA focuses primarily on interworking in the data plane, with some discussion of management plane interworking as well. Control-plane interworking (between different signaling protocols or different routing protocols) is outside the scope of this document.

1.3 Overview

Ethernet multi-service interworking provides ubiquitous Ethernet service between customer sites that use different AC types, so that all traffic (including Layer 3 routing control protocols such as OSPF and IS-IS) is transported transparently between customer sites. This specification applies equally to Virtual Private Wire Service (VPWS), to Virtual Private LAN Service (VPLS) [L2 Frame], and to other Ethernet services [MEF 6]. Ethernet multi-service interworking offers the flexibility of supporting such applications at the expense of the additional overhead associated with Ethernet encapsulation (e.g., Ethernet header and FCS) and configuration changes at the CEs.

Some of the procedures used in IGP routing protocols are dependent upon the underlying Layer 2 protocol. For a point-to-point connection (e.g., an ATM VC) these procedures are different from the procedures for a broadcast connection (e.g., Ethernet).

Section 4 describes different models for Ethernet interworking. Section 5 provides a description of encapsulation mapping of the native Ethernet Service among different attachment circuit types. Section 6 describes encapsulation formats. Section 7 describes traffic management interworking and Section 8 describes PVC management interworking.

The descriptions in Sections 7 and 8, concerning traffic management and PVC management interworking, respectively, utilize the Ethernet Service model defined by the Metro Ethernet Forum (MEF) [MEF6, MEF10]. No attempt is made to define new traffic and service parameters for Ethernet circuits.. At the time of publication, the MEF continues to work on Ethernet Service definitions, and Traffic Management specifications.

2 Definitions and Terminology

2.1 Definitions

Must, Shall or Mandatory — the item is an absolute requirement of this specification.

Should — the item is desirable.

May or Optional — the item is not compulsory, and may be followed or ignored according to the needs of the implementer.

2.2 Acronyms and Abbreviations

Acronym or Abbreviation	Definition
AC	Attachment Circuit
AAL	ATM Adaptation Layer
ABR	Available Bit Rate
ATM	Asynchronous Transfer Mode
AUU	ATM User-to-User indication
Bc	Committed Burst
Be	Excess Burst
BECN	Backward Explicit Congestion Notification
BPDU	Bridge Protocol Data Unit
CBS	Committed Burst Size

Acronym or Abbreviation	Definition
CDVT	Cell Delay Variation Tolerance
CE	Customer Edge
CIR	Committed Information Rate
CLP	Cell Loss Priority
CoS	Ethernet Class of Service
CRC	Cyclic Redundancy Check
DA	Destination Address
DE	Discard Eligibility
DLCI	Data Link Connection Identifier
DSCP	DiffServ Code Point
EBS	Excess Burst Size
EFCI	Explicit Forward Congestion Indicator
EIR	Excess Information Rate
EoPW	Ethernet over PW
ESI	Ethernet Service Instance
EVC	Ethernet Virtual Connection
FECN	Forward Explicit Congestion Notification
FCS	Frame Check Sequence
FR	Frame Relay
GARP	Generic Attribute Registration Protocol
GFR	Guaranteed Frame Rate
GMRP	GARP Multicast Registration Protocol
GVRP	GARP VLAN Registration Protocol
HDLC	High-level Data Link COntnrol
IEEE	Institute of Electrical and Electronics Engineers
IETF	Internet Engineering Task Force
IGP	Interior Gateway Protocol
IP	Internet Protocol
IS-IS	Intermediate System – Intermediate System
ITU-T	International Telecommunication Union – Telecommunication Standardization Sector
IWF	Interworking Function
L2VPN	Layer 2 Virtual Private Network
LACP	Link Aggregation Control Protocol

Acronym or Abbreviation	Definition
LAN	Local Area Network
LCP	Link Control Protocol
LDP	Label Distribution Protocol
LER	Label Edge Router
LLC	Logical Link Control
LR	Link Rate
LSP	Label Switched Path
LSR	Label Switching Router
MAC	Media Access Control
MBS	Maximum Burst Size
MCR	Minimum Cell Rate
MEF	Metro Ethernet Forum
MFS	Maximum Frame Size
MPLS	Multi Protocol Label Switching
NCP	Network Control Protocol
NNI	Network Node Interface
NSP	Native Service Processing
OAM	Operation And Maintenance
OSPF	Open Shortest Path First
OUI	Organizationally Unique Identifier
PCR	Peak Cell Rate
PDU	Protocol Data Unit
PE	Provider Edge
PID	Protocol Identifier
PPP	Point to Point Protocol
PVC	Permanent Virtual Connection
PW	Pseudo wire
PWE3	Pseudo-wire Emulation Edge-to-Edge
QoS	Quality of Service
RFC	Request For Comment
SA	Source Address
SAR	Segmentation and Reassembly
SCR	Sustainable Cell Rate
SDU	Service Data Unit

Acronym or Abbreviation	Definition
SFD	Start-of-Frame Delimiter
STP	Spanning Tree Protocol
SVC	Switched Virtual Connection
TBD	To Be Determined
TM	Traffic Management
ToS	Type of Service
UBR	Unspecified Bit Rate
UNI	User to Network Interface
VC	Virtual Connection
VCI	Virtual Channel Identifier
VLAN	Virtual LAN
VPI	Virtual Path Identifier
VPLS	Virtual Private LAN Service
VPN	Virtual Private Network
VPWS	Virtual Private Wire Service

2.3 Reference

2.3.1 Normative References

The following is a list of standards on which this Specification is based:

[RFC 3985] S. Bryant, P. Pate, Pseudo Wire Emulation Edge-to-Edge (PWE3) Architecture, March 2005

[RFC 2684] D. Grossman, J. Heinanen, Multiprotocol Encapsulation over ATM Adaptation Layer 5, September 1999.

[RFC 2427] C. Brown, A. Malis, Multiprotocol Interconnect over Frame Relay, September 1998.

[RFC 2878] M. Higashiyama, F. Baker, PPP Bridging Control Protocol (BCP), July 2000.

[IEEE 802.1ad] IEEE Standards for Local and metropolitan area networks—Virtual Bridged Local Area Networks – Amendment 4: Provider Bridges, August 17, 2005

[IEEE 802.1D] IEEE Standards for Local and metropolitan area networks—Media Access Control (MAC) Bridges, June 2004

[IEEE 802.1Q] IEEE Standards for Local and metropolitan area networks—Virtual Bridged Local Area Networks, May 2003

[Y.1415] ITU-T Recommendation on Ethernet-MPLS network interworking - User plane interworking, Feb 2005.

[FM-MSInt] MFA Forum Specification on Fault Management for Multiservice Interworking version 1.0, MFA Forum 13.0.0, June 2006

[EthoMPLS] Luca Martini, Encapsulation Methods for Transport of Ethernet over MPLS Networks, Work in progress, November 2005, draft-ietf-pwe3-ethernet-encap-11.txt

2.3.2 Informative References

[L2 Frame] Loa Andersson, Eric C. Rosen, Framework for Layer 2 Virtual Private Networks (L2VPNs), Work in progress, June 2004, draft-ietf-l2vpn-l2-framework-05.txt

[I.363.5] ITU-T Recommendation B-ISDN ATM Adaptation later specification: Type 5 AAL, 08/96.

[FRF.8.2] Frame Relay / ATM PVC Service Interworking Implementation Agreement - February 9, 2004

[ATM B-ICI] ATM Forum BISDN Inter Carrier Interface (B-ICI) Specification, Version 1.0, August 1993

[MEF 6] Metro Ethernet Forum, MEF 6, Ethernet Service Definition, Phase 1, June 2004

[MEF 10] Metro Ethernet Forum, MEF 10, Ethernet Services Attributes, Phase 1, November 2004

[I.370] ITU-T Recommendation I.370, Integrated Services Digital Network (ISDN), Overall Network Aspects And Functions, ISDN User-Network Interfaces, Congestion Management For The ISDN Frame Relaying Bearer Service, 1991.

3 Reference Model

Figure 1 shows an interworking reference model, where CE, PE, pseudo wire (PW), PW Processor (including payload encapsulation, LDP signaling, etc. for PW processing), and the Attachment Circuits (ACs) are defined in IETF PWE3 Working Group terminology and specifications [RFC 3985]. Specifically, the AC is a physical or virtual circuit connection between a CE and a PE. From a PE point of view, the AC can be on either a UNI or an NNI. If the AC is FR or ATM, it must be determined whether the AC is carrying an Ethernet UNI (frames from an Ethernet end-station) or NNI (frames coming from an Ethernet network)

.

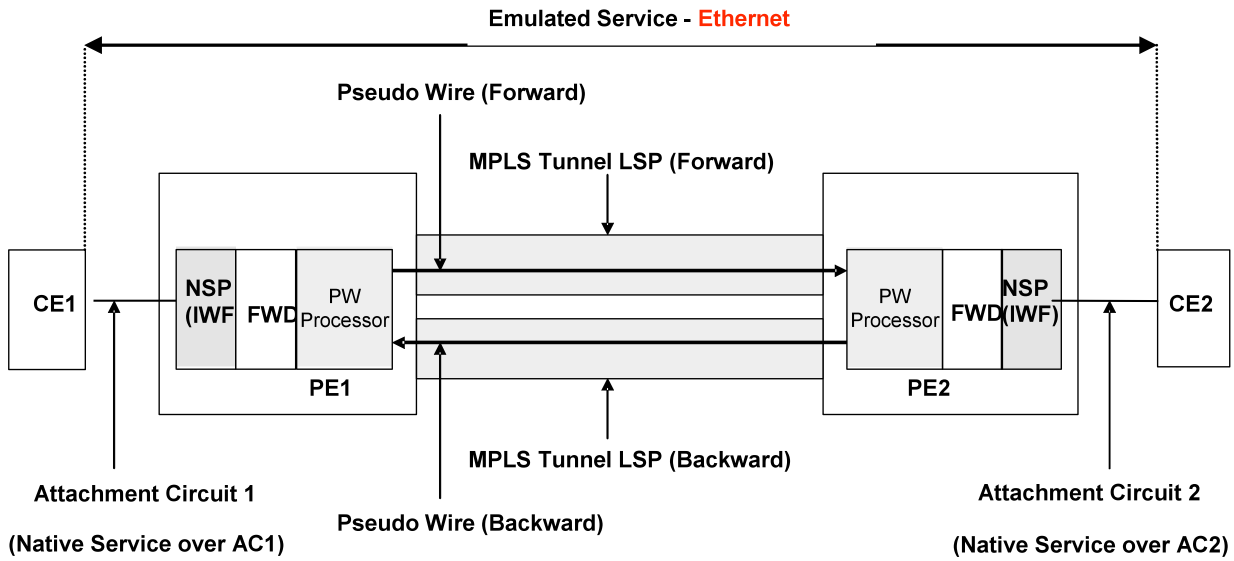

FIGURE 1. Interworking Reference Model

With respect to Multi-service Interworking, it is important to differentiate between an AC and the Native Service that runs over that AC. The AC is a physical or virtual circuit connection between a CE and a PE. However, the Native Service in this reference model is the common Ethernet application service that runs between a pair of CEs and is carried over the two disparate ACs. For example, an AC between a CE and a PE can be ATM or FR, but the Native Service is Ethernet. Sometimes the AC and the Native Service are of the same type (e.g., both are of type Ethernet), in which case it simplifies the interworking task performed by the Native Service Processing (NSP) module on the corresponding PE.

The NSP module is responsible for processing of the data received by the PE (including both data and signaling traffic) from the CE before presentation to the PW for transmission across the core, or processing of the data received from a PW by a PE before it is output on the AC. This document specifies the following interworking functions by the NSP module for Ethernet interworking over MPLS.

- Encapsulation mapping of the native Ethernet Service between ATM/FR/PPP/Ethernet AC and their corresponding PW – e.g., identification and mapping of each service instance for each AC to/from their corresponding PW
- Traffic Management Interworking - Mapping of traffic and service parameters
- PVC Management Interworking - Mapping PVC management indications between the ATM/FR/Ethernet AC and their corresponding PW

The forwarder module (FWD) in Figure 1 is responsible for forwarding the frames between the ACs and the PWs corresponding to a given service instance. There is one such forwarder module per service instance (per customer). In the case of multipoint-to-multipoint connectivity, the Ethernet header (VLAN and MAC address) is used in the forwarding decision. However, in the case of point-to-point connectivity, there is a static mapping between a given AC and its associated PW, and thus the AC identifies which PW is to be used for forwarding the frames to the remote PE (and vise versa - the PW identifies which AC is to be used).

It should be noted that the frame adaptation functionality (formatting the frames based upon the AC type) and filtering/blocking Ethernet control frames (if any) is done by the NSP module (more specifically by the Interworking Function within this module) and not by the forwarder module. It should also be noted that the presentation of the PE in Figure 1 is logical, and that its modules can be implemented by different physical entities.

4 Interworking Models

There are two models for Ethernet Interworking. In the first model the local attachment circuits are terminated at the PE. The Ethernet PDUs are then encapsulated in an Ethernet pseudo wire and transported to the remote PE. From there the PDUs are de-capsulated from the Ethernet pseudo wire and sent on the remote attachment circuit. This model is referred to as the two-sided model, and is described in section 4.1. Implementation conformance to this model is mandatory.

The second model extends the non-Ethernet AC from the local PE to the remote PE by using a pseudo wire of the corresponding AC type. The pseudo wire is terminated at the remote PE, where the Ethernet PDUs are de-capsulated and sent on the remote attachment circuit. This second model, which is referred to as a single-sided model, is described in Annex A. Implementation conformance to this single-sided model is optional.

Note: If one of the attachment circuits is of type Ethernet, then the non-Ethernet AC is extended over the pseudo wire across the MPLS network.

4.1 Two-Sided Model

Figure 2 illustrates the two-sided model for Ethernet Interworking. In this model each PE terminates its Attachment Circuit locally, and after extraction of the Ethernet frames, transports them over a pseudo wire of type Ethernet. If the Attachment Circuit is of type ATM, FR, or PPP, then the AC is terminated locally, and the Ethernet frames are encapsulated and de-capsulated using the procedures outlined in [RFC 2684], [RFC 2427], or [RFC 2878] respectively. If the AC is of type Ethernet, and if the local PE runs the Link Aggregation Control Protocol (LACP - 802.3ad) to the attached CE, then this protocol shall be terminated on the local PE. Other Ethernet protocols such as 802.1x, GVRP, GMRP, STP, etc. may be tunneled, discarded, or processed based upon Ethernet UNI service attributes, as identified in [MEF 6].

The two-sided model has the following characteristics:

1. Each PE is only required to support the procedures for the interfaces that it supports. For example, a PE with an ATM interface is only required to support RFC 2684, and not RFC 2427 or other functionality that is not pertinent to that interface.
2. Each PE is only required to support a common Ethernet pseudo wire irrespective of the Attachment Circuit types on the remote PEs.
3. Configuration of each AC or VC type is limited to the PE that supports that interface.
4. Reduced software overhead - for example, if PPP pseudo wires are extended to the FR side, then the FR PE needs to support PPP NCP/LCP protocols, whereas, if PPP is terminated locally and only the Ethernet payload is transported, then this additional protocol burden is not imposed on the FR PE.

5. Supporting VPLS service (as well as VPWS) – since each AC is terminated locally at the corresponding PE, then each PE has access to Ethernet headers for making a forwarding decision (based upon destination MAC addresses and VLANs)

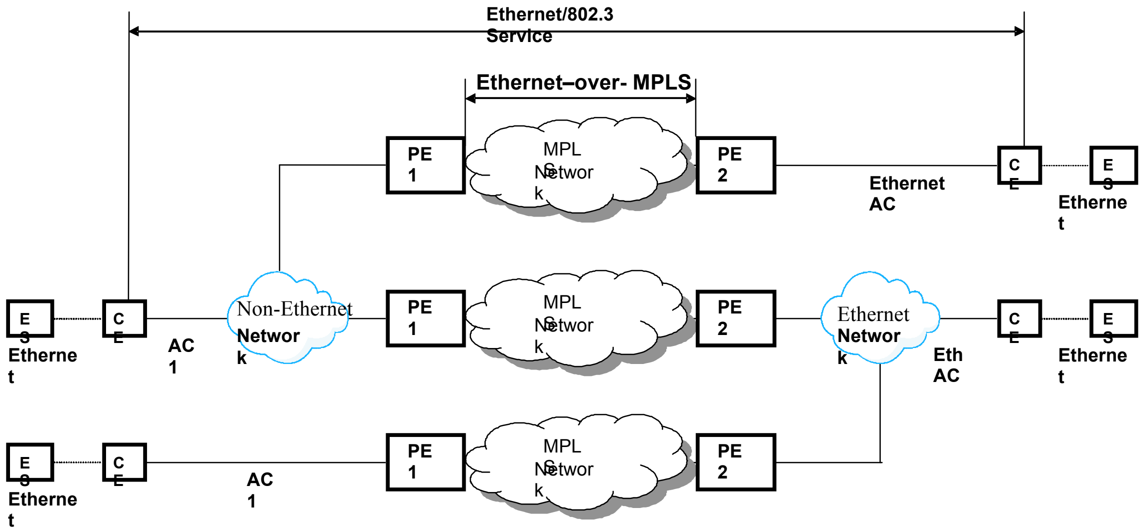

FIGURE 2. General Model for Ethernet Interworking

Table 1 lists all possible combinations of ACs that can exist for this model, and the applicable RFC procedures that need to be performed by the IWF.

	AC – 1	AC – 2	Encap. in PE1	Encap. in PE2
1	Ethernet	ATM	N/A	RFC 2684-B
2	Ethernet	FR	N/A	RFC 2427-B
3	Ethernet	PPP/HDLC	N/A	RFC 2878
4	FR	ATM	RFC 2427-B	RFC 2684-B
5	FR	PPP/HDLC	RFC 2427-B	RFC 2878
6	ATM	PPP/HDLC	RFC 2684-B	RFC 2878

Table 1. Encapsulation methods for Ethernet Interworking

4.2 CE Configuration Changes

For Ethernet Interworking the CE routers or hosts with non-Ethernet ACs need to be configured for Ethernet encapsulation. There are currently two common ways of configuring the CE routers for such operation; they are:

- Configure the CE router interface as an IEEE bridged interface
- Configure the CE router interface as a routed interface with bridged encapsulation

The first method requires that the CE router interface be configured for bridging operation, and be connected to routed interfaces by way of a virtual interface. Routed packets for a given protocol can be exchanged between routed and bridged interfaces via this virtual interface. This virtual interface acts like a normal routed interface; it represents the corresponding bridge group to routed interfaces within the CE router.

The second method requires that the CE interface be configured for routed operation, but it has to provide bridged encapsulation via the interface. Using this method, the interface looks and behaves like a routed interface to all other routed interfaces, although packets that are sent out of this interface are encapsulated with an Ethernet header. In the case of an ATM/FR routed interface with bridged encapsulation, this would require a specific IGP configuration to force the router to see the interface as a broadcast interface.

5 Ethernet Service Instance & Class of Service

5.1 Ethernet Service Instance - ESI

The MEF defines an Ethernet Virtual Connection (EVC) as an association of two or more UNIs, where the UNI is a standard Ethernet interface and point of demarcation between Customer Equipment and a service provider's network. An EVC is either point-to-point or multipoint-to-multipoint. For example, a customer with five sites (e.g., five pieces of Customer Equipment - CEs) can be connected via a single multipoint-to-multipoint EVC, or the sites can be connected via ten point-to-point EVCs in a full-mesh configuration, with each CE having four point-to-point EVCs to the other four CEs.

We define an Ethernet Service Instance (ESI) as an association of two or more Attachment Circuits (that can be UNIs and/or NNIs) over which Ethernet service is offered to a given customer. An AC can be an Ethernet interface (as with an EVC UNI), or it can be an ATM or FR VC or PPP/HDLC interface.

An ESI corresponds to a VPLS instance (for a multipoint connection) or to a VPWS instance (for a point-to-point connection) as defined in [L2-Frame]. Furthermore, in the context of Provider Bridges [IEEE 802.1ad] and multipoint connections, an ESI corresponds to a Service VLAN (S-VLAN) as defined by [IEEE-802.1ad].

It should be noted that an ESI cannot be directly compared with an EVC per MEF definition since an ESI is associated with a set of ACs, whereas an EVC is associated with a set of UNIs. However, if one limits the ACs associated with a given ESI to only UNIs, then for a multipoint connection an ESI corresponds to a multipoint-to-multipoint EVC, and for a point-to-point connection; an ESI corresponds to a number of point-to-point EVCs where the number of EVCs is proportional to the square of the number of associated UNIs for that service instance.

An ESI (either for point-to-point or multipoint connectivity) is associated with a forwarding path within the service provider's network; it is different from an Ethernet Class of Service (CoS – described in section 5.2) which is associated with the frame QoS treatment by each node along the path defined by the ESI. An ESI can have one or more CoS associated with it.

The following table lists different mappings of customer Ethernet traffic at an AC to its corresponding service instance. As can be seen, there are several possible ways to perform such mapping. In the first scenario it is assumed that a physical port (Ethernet/PPP) or a logical port (ATM/FR VC) only carries untagged traffic, and all the traffic is mapped to the corresponding service instance (ESI); this is referred to as “port-based w/ untagged traffic”. In the second scenario it is assumed that a physical or a logical port carries both tagged and untagged traffic, and all that traffic is mapped to the corresponding service instance (ESI); this is referred to as “port-based w/ tagged and untagged traffic”. In the third scenario it is assumed that only a single VLAN from the Ethernet physical interface is mapped to the corresponding service instance (ESI); this is referred to as “VLAN mapping”. Finally, in the fourth scenario it is assumed that a group of VLANs from the Ethernet physical interface is mapped to the corresponding service instance (ESI); this is referred to as “VLAN bundling”. The application of the third and fourth scenarios to ATM/FR virtual circuits and/or PPP/HDLC interfaces is not required, as shown in the table; therefore, the only applicable interface type for the third and fourth scenarios is Ethernet.

	Interface Type		
Mapping Modes at an Attachment Circuit (for a given service instance)	**Ethernet Interface**	**ATM/FR VC**	**PPP/HDLC Interface**
Port-based - untagged only	S	S	S
Port-based - tagged & untagged	S	S	S
VLAN mapping	S	NS	NS
VLAN bundling	S	NS	NS

S Supported for the Interface type

NS Not specified in this version

Note: The table provides mapping modes for each interface type. Equipment may support a subset of the modes applicable to a given interface type.

Table 2. Mappings Modes Supported for Different Interface Types

It should be noted that for a given service instance (ESI) the mapping does not need to be uniform across all its ACs even though the majority of deployment scenarios require uniform mapping. For example, a service that consists of several ATM VCs (one VC per branch office) and a single Ethernet physical interface (for the headquarters office), can have “port-based w/ untagged traffic” for all its ATM VCs and can have “VLAN mapping” for its Ethernet interface. In this example, the untagged traffic from each ATM VC is mapped to a unique VLAN over the Ethernet interface. The service in this example provides point-to-point ESIs between the headquarters and each branch office. It should be noted that not all possible mapping combinations are allowed for a given service instance. For example, it is not possible to have a mapping between “port-based w/ untagged traffic” and “VLAN bundling”. Tables 3 and 4 show the possible combinations allowed for traffic mapping among Ethernet interfaces, as well as between Ethernet and ATM/FR/PPP interfaces for a given service instance.

It should also be noted that the possible mapping combinations listed in Table 3 are more than the ones defined and allowed by the MEF. The reason for this is that these mappings are defined for Attachment Circuits (ACs) which can be either an Ethernet UNI or an Ethernet NNI, and in the case of an Ethernet NNI, the Ethernet frame format based upon [IEEE 802.1ad] can exist. In the MEF, however, the mapping is only defined for UNIs where Ethernet frame formats are based on [IEEE 802.1D] or [IEEE 802.1Q].

		Ethernet I/F & Associated Service Instance(s)			
		Port-based w/ untagged	Port-based w/ tagged & untagged	VLAN mapping	VLAN bundling
Ethernet I/F	Port-based w/ untagged traffic	A	NS	A	N/A
	Port-based w/ tagged & untagged	NS	A	A (Note-1)	A (Note-2)
	VLAN mapping	A	A (Note-1)	A	A (Note-3)
	VLAN bundling	N/A	A (Note-2)	A (Note-3)	A

A Mapping between the two modes are Allowed NS Not specified in this version
N/A Not Allowed

Table 3. Different Mappings among Ethernet I/Fs Associated with a Service Instance

Note - 1: In this asymmetric mapping scenario, it is assumed that the CE device with a "VLAN mapping" AC is a Provider Bridge, as defined by [IEEE 802.1ad], because it will receive Ethernet frames with two tags; the outer tag is S-VLAN and the inner tag is C-VLAN received from "port-based" AC.

Note – 2: In this asymmetric mapping scenario, it is assumed that the VLAN tags in the port-based AC are identical to the VLAN tags in the "VLAN bundle" AC, and the PE passes them to the CE device transparently without any processing. Furthermore, it is assumed that the untagged frames (if any) from the port-based UNI are dropped, unless the PE on the "VLAN bundle" AC side has the capability to process a single untagged control stream and branch it into multiple streams (one per bundle), as described in [IEEE 802.1ad].

Note – 3: In this asymmetric mapping scenario, it is assumed that the CE device with a "VLAN mapping" AC is a Provider Bridge, as defined by [IEEE 802.1ad], because it will receive Ethernet frames with two tags; the outer tag is S-VLAN and the inner tag is C-VLAN, received from a "VLAN bundling" AC.

If a PE uses a S-VLAN tag for a given ESI (either by adding a S-VLAN tag to customer traffic or by replacing a C-VLAN tag with a S-VLAN tag), then the frame format and EtherType for the S-VLAN shall adhere to [IEEE 802.1ad].

When a port-based or a VLAN-bundling traffic mapping is used at an AC, and if the PE uses an additional S-VLAN tag to mark the customer traffic received over that AC as belonging to a given ESI, then that PE shall strip the S-VLAN tag before sending the customer frames over the same AC. However, when VLAN-mapping mode is used at an AC, and if the PE uses a S-VLAN tag locally, then there are two modes of operation for customer traffic mapping on that AC, depending upon the AC type (whether it is a UNI or NNI). If the Ethernet interface is a UNI, then the tagged frames over this interface shall have a frame format based upon [IEEE 802.1q], and the PE shall map the provider tag (S-VLAN) back to the customer tag (C-VLAN) before sending it over that AC. However, if the Ethernet interface is an NNI, then the tagged frames over this interface shall have a frame format based upon [IEEE 802.1ad], and the PE may need to translate the tagged frames from one S-VLAN tag into another S-VLAN tag over this AC.

Table 4 lists possible mapping of customer Ethernet traffic that can exist between Ethernet and non-Ethernet interfaces.

		ATM/FR VC or PPP/HDLC I/F & Associated Service Instance(s)			
		Port-based w/ untagged	Port-based w/ tagged & untagged	VLAN mapping	VLAN bundling
Ethernet I/F	Port-based w/ untagged traffic	A	NS	NS	NS
	Port-based w/ tagged & untagged	NS	A	NS	NS
	VLAN mapping	A	A (Note-1)	NS	NS
	VLAN bundling	N/A	A (Note-2)	NS	NS

A Mapping between the two modes are Allowed N/A Not Allowed

NS Not specified in this version

Table 4. Different Mappings of Ethernet I/Fs and ATM/FR VCs Associated with a Service Instance

Note - 1: See Note 1 of Table 3.

Note – 2: See Note 2 of Table 3.

5.2 Ethernet Class of Service - CoS

Ethernet Class of Service (CoS) refers to the Quality of Service (QoS) treatment of all Ethernet frames that belong to that class. The QoS treatment includes, but is not limited to, policing, marking, queuing, and shaping of the Ethernet frames according to a set of traffic parameters. Ethernet Traffic Management will be described in detail in Section 7. Different Layer 2 technologies have different ways of indicating the CoS for their traffic. For example, in the case of ATM, the CoS is associated implicitly with the VC (VPI/VCI), whereas, in the case of Ethernet, in addition to implicitly associating a class of service with a VLAN, there is also a CoS field that can explicitly indicate the class of service to which the received frame belongs.

As mentioned before, there is a main distinction between Ethernet Service Instance (ESI) and Ethernet Class of Service (CoS). The former refers to forwarding treatment of the frames (e.g., which forwarding path is to be chosen/taken within the service provider's network); whereas, the latter refers to the QoS treatment of the frames (e.g., how the frames should be treated on each node along the selected path). An ESI can have one or more CoS associated with it. For example, in the case of the "Port-based (tagged & untagged)" scenario for an Ethernet AC: associating it with an ESI using a provider S-VLAN tag, the options are a) assign a single CoS to that ESI regardless of the CoS field of the customer tags (C-VLANs), or b) have several CoS for that ESI based upon the CoS field of the customer tags (C-VLANs).

If, besides having one or more CoS per ESI, there is a requirement to have a CoS for the aggregate ESIs going through a given AC, then a hierarchical traffic manager (e.g., hierarchical policer or shaper) is needed. The hierarchical traffic manager is for future study. When Ethernet frames are delivered over an ATM or a FR VC to a CE, then there is already a CoS associated with that ATM or FR VC. If, besides having the CoS associated with the VC, there is a requirement to process the CoS field of the Ethernet frames, then that would require a hierarchical traffic manager which is again for future study. Therefore, in the case of non-Ethernet interfaces (e.g., PPP/HDLC, ATM/FR VC), the CoS options for now are limited to the following two choices:

- use the CoS associated with the ATM/FR VC and ignore the CoS field of Ethernet frames, or
- use the CoS field of the Ethernet frames and ignore the CoS associated with ATM/FR VC.

The latter one allows for multiple CoS per VC.

CoS selection at the ACs belonging to a given ESI shall be uniform. In other words, if CoS is selected based upon circuit-id on a given AC, it must also be selected based upon circuit-id for all other ACs associated with that ESI. Furthermore, if CoS is selected based upon IEEE 802 .1p bits or DSCP bits on a given AC, then it shall be done in the same way for other ACs associated with that ESI – e.g., there shall be a single CoS option per ESI.

Tables 5 and 6 shows the allowable CoS selection for different types of mapping on Ethernet and non-Ethernet UNI. The tables also have a column for CoS selection based upon the IP DSCP/ToS field. Even though the DSCP/ToS field is part of the Layer 3 IP header, it still can be used as a CoS identifier since it does not affect the forwarding decision which is determined by the ESI, as identified previously.

The DSCP-based option is only applicable to IP traffic, and thus when this option is chosen for a given ESI, it is assumed that the traffic associated with that ESI and all its corresponding ACs is of type IP. If non-IP packets (with Ethernet encapsulation) are received over any AC associated with the DSCP-based CoS option, then these packets shall get forwarded as Best Effort traffic without any QoS guarantee.

	Class of Service Selection for Ethernet I/Fs		
Mapping at Ethernet Attachment Circuit	**Fixed per ESI**	**CoS (.1p) based per ESI**	**DSCP-based per ESI**
Port-based w/ untagged traffic	A	N/A	A
Port-based w/ tagged & untagged	A	A	A
VLAN mapping	A	A	A
VLAN bundling	A	A	A

A Option Allowed N/A Not Allowed

Table 5. CoS Selection for different AC mappings on Ethernet Interfaces

As noted earlier, there shall be only a single CoS option (fixed, CoS-based, or DSCP-based) per ESI. However, in certain applications, besides having an individual CoS option per ESI, there may be a requirement to apply the fixed-CoS option per group of ESIs (e.g., per aggregate ESI - physical/logical interface carrying the ESIs). Such applications would require hierarchical traffic management, which is for future study, as stated previously.

	Class of Service Selection for ATM/FR VCs or PPP/HDLC I/Fs		
Mapping at non-Ethernet Attachment Circuit	**Fixed per ESI**	**CoS (.1p) based per ESI**	**DSCP-based per ESI**
Port-based w/ untagged traffic	A	N/A	A
Port-based w/ tagged & untagged	A	A	A
VLAN mapping	NS	NS	NS
VLAN bundling	NS	NS	NS

A Option Allowed N/A Not Allowed
NS Not specified in this version

Table 6. CoS Selection for different AC mapping on non-Ethernet Interfaces

5.3 Ethernet Pseudo Wire Types

The IETF PWE3 Working Group has defined two modes of operation for Ethernet pseudo wires [EthoMPLS or Y.1415]: raw mode versus tagged mode. In tagged mode, each Ethernet frame MUST contain a VLAN tag and the tag must be a service-delimiting tag (e.g., to identify a service instance). In raw mode, a frame MAY contain a VLAN tag and if it does, the tag is not a service delimiting tag, and it passes through the PE transparently.

Both raw and tagged modes can be used for Ethernet service interworking with any of the AC traffic mappings described before. If tagged mode EoPW is used, and no service tag (S-VLAN) is present, then the ingress PE must first add a service tag (S-VLAN) prior to sending the frames over the PW. However, if raw mode is used, the ingress PE can send the customer's frames (either tagged or untagged) as is, over the PW without adding/replacing any tag. It should be noted that tagging and untagging customer frames with S-VLAN (service-delimiting tag) is a function local to that PE, and it may or may not be performed by that PE. If a service-delimiting tag (S-VLAN) is used to identify a given service instance, then multiple service instances can share the same pseudo wire. However, if S-VLAN is not used, then there needs to be a separate pseudo wire associated with each service instance.

Table 7 lists the applicability of customer tag (C-VLAN), service tag (S-VLAN), raw mode EoPW, and tagged mode EoPW for each possible mapping at the AC. The C-VLAN and S-VLAN columns show what is allowed with the Ethernet frame at the NSP. The EoPW columns show how the frame is encapsulated.

AC Traffic Mapping	**C-VLAN**	**S-VLAN**	**Raw EoPW**	**Tagged EoPW**
Port-based w/ untagged traffic	N/A	A	A	A
Port-based w/ tagged & untagged	A	A	A	A
VLAN mapping	A	A	A	A
VLAN bundling	A	A	A	A

A Option Allowed N/A Not Allowed

Table 7. C-VLAN, S-VLAN, Raw EoPW, and Tagged EoPW Applicability

As shown in Table 7: for the "port-based w/ untagged traffic" scenario, there is no customer tag (C-VLAN). However, the service provider PE can add a service tag (S-VLAN) to the frames received from that port (Ethernet/PPP I/F or ATM/FR VC) for association of the traffic to a service instance (if a PE is an IEEE 802.1ad Provider Bridge, then the service tag gets added). Now, when the PE wants to send the traffic to the remote PE over the EoPW, it has several choices. If the frame is not tagged with the S-VLAN, then the PE uses raw-mode EoPW; however, if the frame is tagged with the S-VLAN, then it can either strip the S-VLAN tag and send the resultant untagged frame using raw-mode EoPW, or it can send the frame with the S-VLAN tag and use the tagged-mode EoPW.

In the case of the "port-based w/ tagged & untagged" scenario, customer traffic received from the CE can include a customer tag (C-VLAN). With the "port-based untagged" scenario, the PE can add a service tag (S-VLAN) to the receiving frames for its own local processing (e.g., if a PE is an IEE 802.1ad provider bridge, then the S-VLAN is added). When the S-VLAN is added, the frame may contain two tags (one for C-VLAN and the other for S-VLAN). As before, the PE has several options for transmitting this frame over EoPW to the egress PE. If the frame is not tagged with S-VLAN, the PE uses raw-mode EoPW; however, if the frame is tagged with the S-VLAN, then the PE can either strip the S-VLAN tag and send the resultant customer frame (that may contain a C-VLAN tag) using raw-mode EoPW, or the PE can send the frame with the S-VLAN tag and use the tagged-mode EoPW. It should be noted that when the raw-mode EoPW is used, the customer frames can contain C-VLAN tags but these tags are passed transparently to the CE since they are not used as service-delimiting tags.

6 Ethernet Service Interworking – Encapsulation Format

This section illustrates encapsulation frame formats for the following interworking scenarios:

- Frame Relay to Ethernet
- ATM to Ethernet
- ATM to FR
- PPP to Ethernet

6.1 Frame Relay <-> Ethernet Interworking

Figure 3 illustrates the format for encapsulation/de-capsulation of the Ethernet frame by the interworking module at the Ethernet and FR PEs. As described in section 5.1, only port-based mapping for FR VC is required (e.g., all Ethernet traffic coming over the FR VC is associated with a single service instance and is mapped into the Ethernet PW; no VLAN-mapping or VLAN-bundling at the FR AC is mandated).

The Ethernet PW can either be in raw-mode or tagged-mode, depending upon whether a service-delimiting tag needs to be carried over the EoPW or not (tagged-mode must be used if a service-delimiting tag is carried). If tagged-mode is used, then the IWF will add a service-delimiting tag (S-VLAN) to the customer Ethernet frames, which can be either tagged or untagged. Regardless of whether the received customer frames are tagged or untagged, after addition of a S-VLAN tag to the customer frames by the PE, the resultant frame format and EtherType shall adhere to [IEEE 802.1ad].

If the PW is of tagged-mode, then upon receiving the Ethernet frames over the EoPW, the IWF in the Ethernet PE may or may not remove the service-delimiting tag depending upon the type of the mapping at the Ethernet AC. If the mapping is of type port-based or VLAN-bundling, then the service-delimiting tags will be removed before the Ethernet frames are sent over the Ethernet AC. If the mapping is of type VLAN-mapping, then the service-delimiting tag (S-VLAN) will either get translated into the customer tag (C-VLAN) or it will get passed as is over the AC. If the AC is of type UNI, then the S-VLAN is translated to C-VLAN prior to frame transmission over the AC (frame format is translated from IEEE 802.1ad to IEEE 802.1q] but if the AC is of type NNI, then a S-VLAN tag is either transmitted over the AC as is, or mapped onto another S-VLAN tag (with IEEE 802.1ad frame format].

Ethernet encapsulation over a FR VC is based upon [RFC-2427]. This specification only supports LLC encapsulation of the bridged protocols with PID values 0x07 and 0x0E. It should be noted that the BPDU frames are treated differently at the FR PE and have a different FR encapsulation format from the one shown in the figure – e.g., the PID is 0x0E and the encapsulation does not contain MAC addresses.

Therefore, upon receiving a BPDU frame over the FR VC, the IWF at the FR PE shall do the following before forwarding the BPDU frame to the egress PE over the EoPW as an LLC Type 1 frame:

1. add a BPDU multicast MAC address (i.e., 0x01-80-C2-00-00-00 per Table 7-10 IEEE 802.1D 2004) as the MAC DA
2. add a dummy unicast MAC address (e.g., MAC address of the IW unit) as the MAC SA.
3. add length field and
4. add the LLC control information of 3 bytes 0x424203 per 802.1D section 7.12.3

In the opposite direction, upon receiving a BPDU frame over EoPW, the IWF at the FR PE shall remove the MAC header, length field and LLC control information, and shall encapsulate the frame based upon [RFC-2427] before delivering it over the FR VC.

BPDU frames may be sent as regular Ethernet frames with PID 0x07 without the MAC SA and DA being removed. In this case no change is required to the Ethernet frame format.

Figures 3 to 6 depict Ethernet encapsulation without FCS retention for some of the possible traffic mappings between an Ethernet AC and a FR AC. This non-FCS retention mode of operation is the only mode supported by the IWF of both PEs. In this mode, the Ethernet PE strips the Ethernet FCS before sending the frame over the PW. During setup of the PW the default is to not use FCS retention. Therefore, upon receiving Ethernet frames over the PW, the FR PE knows that it shall use the encapsulation that does not contain a FCS with a PID of 0x07.

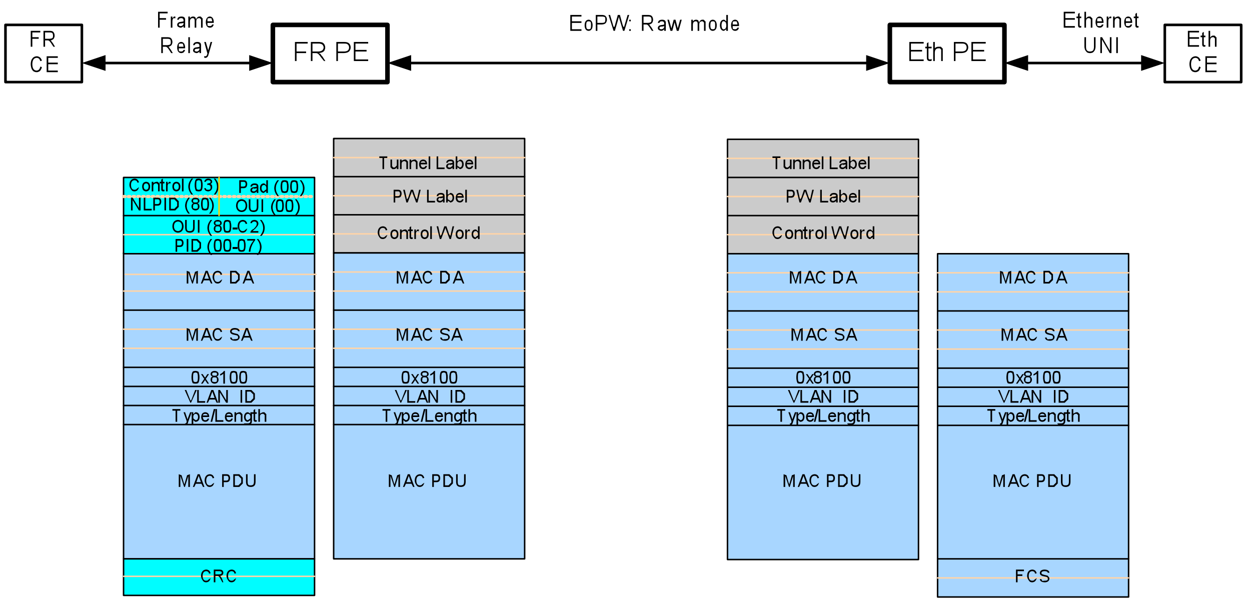

FIGURE 3. Frame Format for FR to Ethernet Interworking without FCS retention – Port-mode traffic mapping for both FR & Ethernet AC w/ Raw mode EoPW

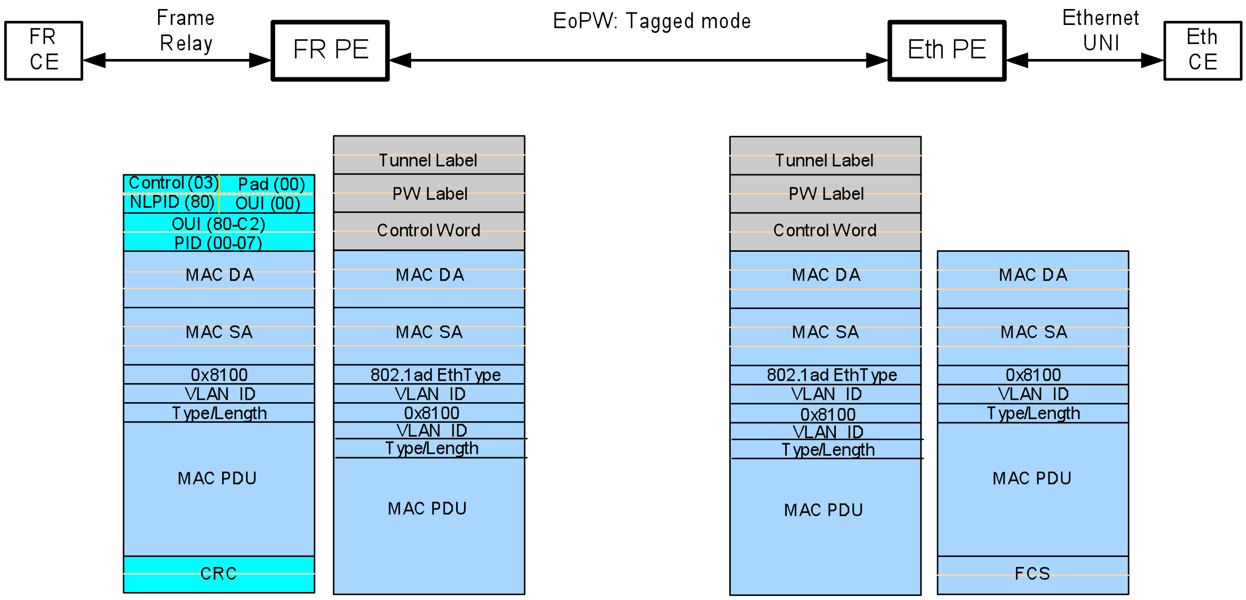

FIGURE 4. Frame Format for FR to Ethernet Interworking without FCS retention – Port-mode traffic mapping for both FR & Ethernet AC w/ Tagged mode EoPW

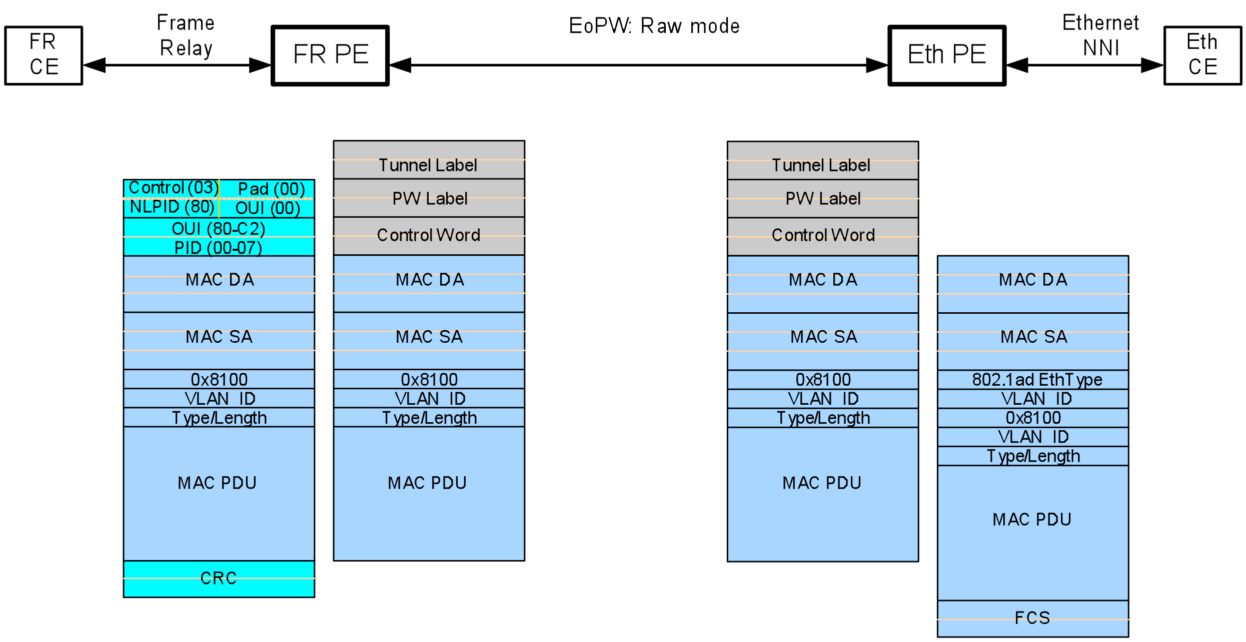

FIGURE 5. Frame Format for FR to Ethernet Interworking without FCS retention – Port-mode traffic mapping at FR & VLAN-mapping mode at Ethernet AC w/ Raw mode EoPW

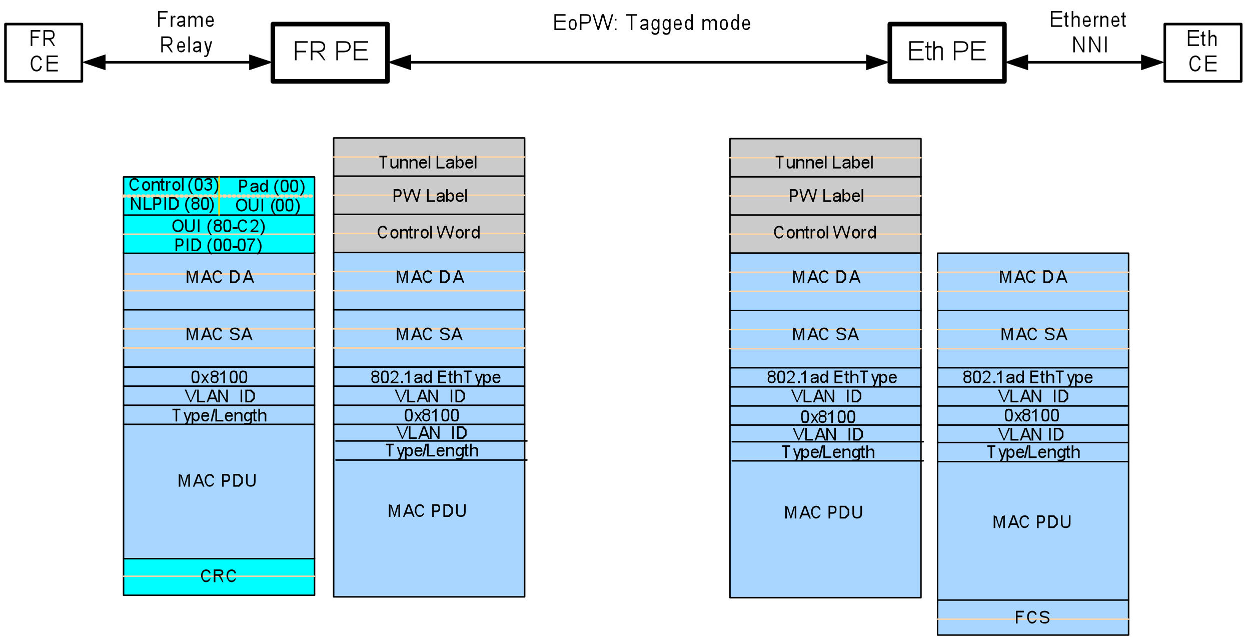

FIGURE 6. Frame Format for FR to Ethernet Interworking without FCS retention – Port-mode traffic mapping at FR & VLAN mode at Ethernet AC w/ Tagged mode EoPW

6.1.1 Frame Formatting and Delimiting

PE connected to the Ethernet CE:

It operates the same way as in the case of Ethernet over MPLS for like-to-like service [EthoMPLS].

PE connected to the Frame Relay CE:

- FR to Ethernet direction: the FR VC is terminated, and the Ethernet MAC frame is extracted from the Frame Relay Bridged frame. Frame flags, inserted zero bits and the FR CRC-16 are stripped, and the Q.922 core frame header is removed. Ethernet encapsulation is based upon [RFC-2427].
- Ethernet to FR direction: the Ethernet MAC frame is encapsulated into the FR bridged encapsulation frame format, adding the CRC-16 and Q.922 core frame header. The FR DLCI is configured to correspond to all Ethernet frames received over the PW (e.g., port-mode).

  The congestion indication fields (BECN, FECN) and Command/Response field (C/R) are set to zero. The association between the FR DLCI and Ethernet Service Instance (ESI) is made at the time the PVC is provisioned.

The BPDUs frames are treated differently at the FR PE, and have a different FR encapsulation format (See Section 6.1 for details).

6.2 ATM to Ethernet Interworking

Figure 7 illustrates the format for encapsulation/de-capsulation of the Ethernet frame by the interworking module at the Ethernet and ATM PEs. As described in section 5.1, only port-based mapping for an ATM VC is required (e.g., all Ethernet traffic coming over the ATM VC is associated with a single service instance and is mapped into the Ethernet PW - no VLAN-mapping or VLAN-bundling at the ATM AC is mandated).

The Ethernet PW can either be in raw-mode or tagged-mode, depending upon whether a service-delimiting tag needs to be carried over the EoPW or not (tagged-mode must be used if a service-delimiting tag is carried). If tagged-mode is used, then the IWF will add a service-delimiting tag (S-VLAN) to the Customer Ethernet frames, which can be either tagged or untagged. Regardless of whether the received customer frames are tagged or untagged, after addition of a S-VLAN tag to the customer frames by the PE, the resultant frame format and EtherType shall adhere to [IEEE 802.1ad].

If the PW is of type tagged-mode, then upon receiving the Ethernet frames over the EoPW, the IWF in the Ethernet PE may or may not remove the service-delimiting tag depending upon the type of mapping at the Ethernet AC. If the mapping is of type port-based or VLAN-bundling, then the service-delimiting tags will be removed before the Ethernet frames are sent over the Ethernet AC. If the mapping is of type VLAN-mapping, then the service-delimiting tag (S-VLAN) will either get translated into the customer tag (C-VLAN) or it will get passed as is over the AC. If the AC is of type UNI, then the S-VLAN is translated to C-VLAN prior to frame transmission over the AC (frame format is translated from IEEE 802.1ad to IEEE 802.1q], but if the AC is of type NNI, then the S-VLAN tag is either transmitted over the AC as is or is mapped onto another S-VLAN tag (with IEEE 802.1ad frame format].

Ethernet encapsulation over ATM is based upon [RFC-2684]. This specification only supports LLC encapsulation of the bridged protocols with PID values 0x07 and 0x0E. It should be noted that the BPDU frames are treated differently at the ATM PE and have a different ATM encapsulation format from the one shown in the figure – e.g., the PID is 0x0E and the encapsulation does not contain a MAC addresses. Therefore, upon receiving a BPDU frame over the ATM VC, the IWF at the ATM PE, shall do the following before forwarding the BPDU frame to the egress PE over the EoPW as an LLC Type 1 frame:

5. add a BPDU multicast MAC address (i.e., 0x01-80-C2-00-00-00 per Table 7-10 IEEE 802.1D 2004) as the MAC DA
6. add a dummy unicast MAC address (e.g., MAC address of the IW unit) as the MAC SA.
7. add length field and
8. add the LLC control information of 3 bytes 0x424203 per 802.1D section 7.12.3

In the opposite direction, upon receiving a BPDU frame over EoPW, the IWF at the ATM PE shall remove the MAC header, length field and LLC control information and shall encapsulate the frame based upon [RFC-2684] before delivering it over the ATM VC.

BPDU frames may be sent as a regular Ethernet frames with PID 0x07 without the MAC SA and DA being removed. In this case there is no change required to the Ethernet frame format.

Figures 7 & 8 depict Ethernet encapsulation without FCS retention for some of the possible traffic mappings between an Ethernet AC and an ATM AC. This non-FCS retention mode of operation is the only mode supported by the IWF of both PEs. In this mode, the Ethernet PE strips the Ethernet FCS before sending the frame over the PW. During setup of the PW the default is to not use FCS retention. Therefore, upon receiving Ethernet frames over the PW, the ATM PE knows that it shall use the encapsulation that does not contain a FCS with a PID of 0x07.

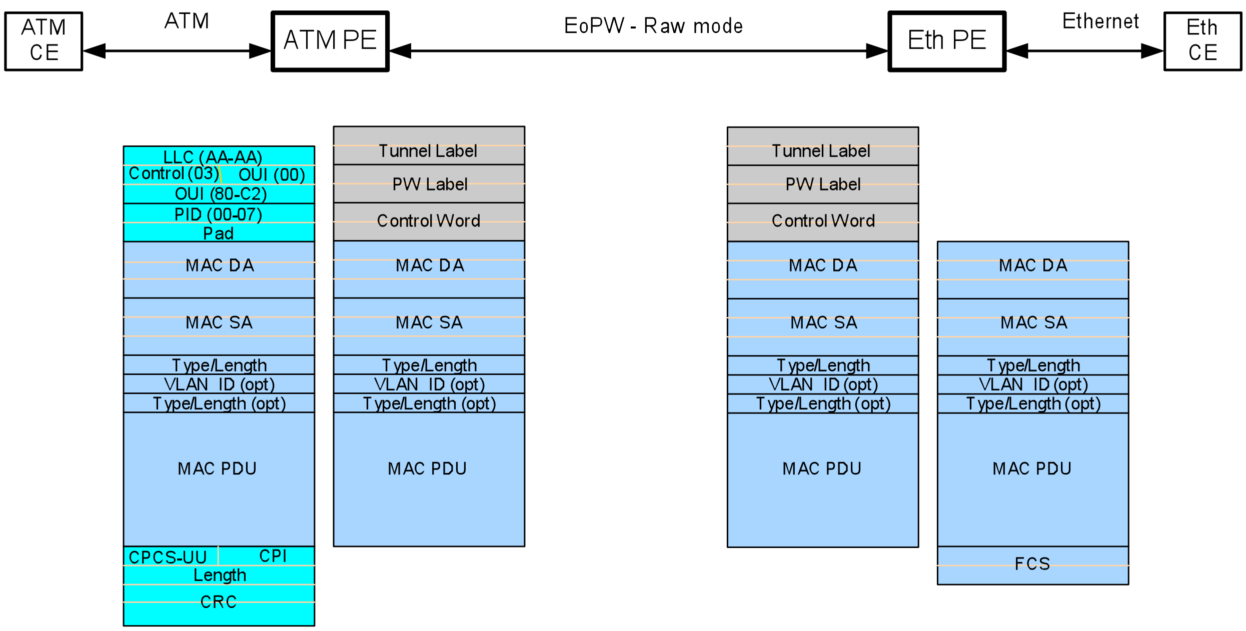

FIGURE 7. Frame Format for ATM to Ethernet Interworking without FCS retention – w Raw mode EoPW

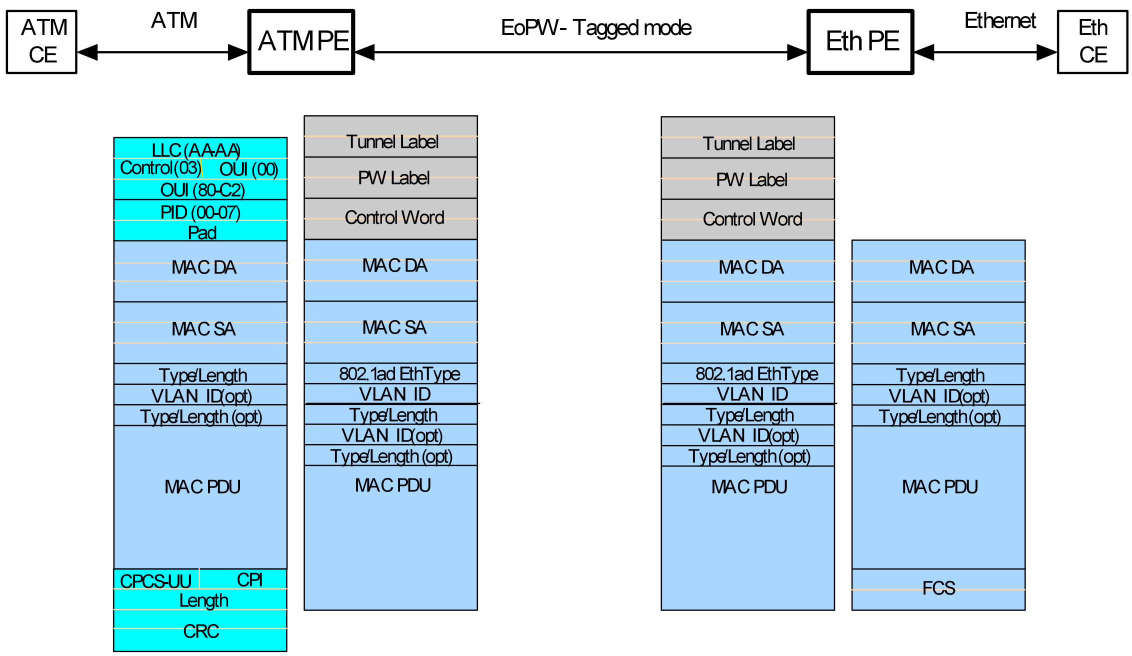

FIGURE 8. Frame Format for ATM to Ethernet Interworking without FCS retention – w/Tagged-mode EoPW

6.2.1 Frame Formatting and Delimiting

PE connected to the Ethernet CE:

It operates in the same way as in the case of Ethernet over MPLS for like-to-like service (See [EthoMPLS]).

PE connected to the ATM CE:

- ATM to Ethernet direction: the Ethernet MAC frame is extracted from the ATM Bridged frame. The message delineation provided by AAL5 is used to identify frame boundaries. The ATM AAL5 SDU identifies the Ethernet MAC frame. Since MAC frames are included in the ATM AAL5 frames with bridged encapsulation, the ATM connection is terminated and the Ethernet MAC frame is delivered based upon the encapsulation in [RFC 2684].

  The Ethernet MAC frame received (with or without FCS) in the ATM AAL5 frame is passed to the Ethernet PW“NSP” function (See [EthoMPLS] for details).

- Ethernet to ATM direction: the Ethernet MAC frame is encapsulated in the ATM bridged encapsulation frame format. The Ethernet Mac frame is passed as AAL type 5 SDU. ITU-T Rec. I.363.5 [I.363.5] is used for reassembly of the AAL5 PDU.

The ATM VPI/VCI is configured to correspond to an Ethernet Service Instance (ESI) – e.g., either VLAN-ID (in Ethernet VLAN-mode) or correspond to all Ethernet frames (in Ethernet port-mode). The association is made at the time the PVC is provisioned.

The congestion indication field (EFCI) and ATM User-to-user indication (AUU) bits are set to zero.

Note: Loss priority parameter mapping is to be provided.

The BPDU frames are treated differently at the ATM PE and have a different ATM encapsulation format (See Section 6.2 for details).

6.3 FR to ATM Ethernet Encapsulation Interworking

When the supported service over FR and ATM layers is Ethernet service, then the ATM and FR links are terminated locally by their corresponding PEs, and Ethernet Encapsulation is used. After extraction of the Ethernet frames, the ingress PE transports them over the network using EoMPLS on the pseudo wire. The egress PE, after performing label disposition, encapsulates the Ethernet frames using the Ethernet encapsulation format based upon [RFC 2427] for the FR CE, and based upon [RFC 2684] for the ATM CE.

Details of the format for encapsulation/de-capsulation of the Ethernet frame by the interworking module at the FR PE are in Section 6.1.

Details of the format for encapsulation/de-capsulation of the Ethernet frame by the interworking module at the ATM PE are in Section 6.2.

Figures 9 depict Ethernet encapsulation without FCS retention for some of the possible traffic mappings between an ATM AC and an FR AC.

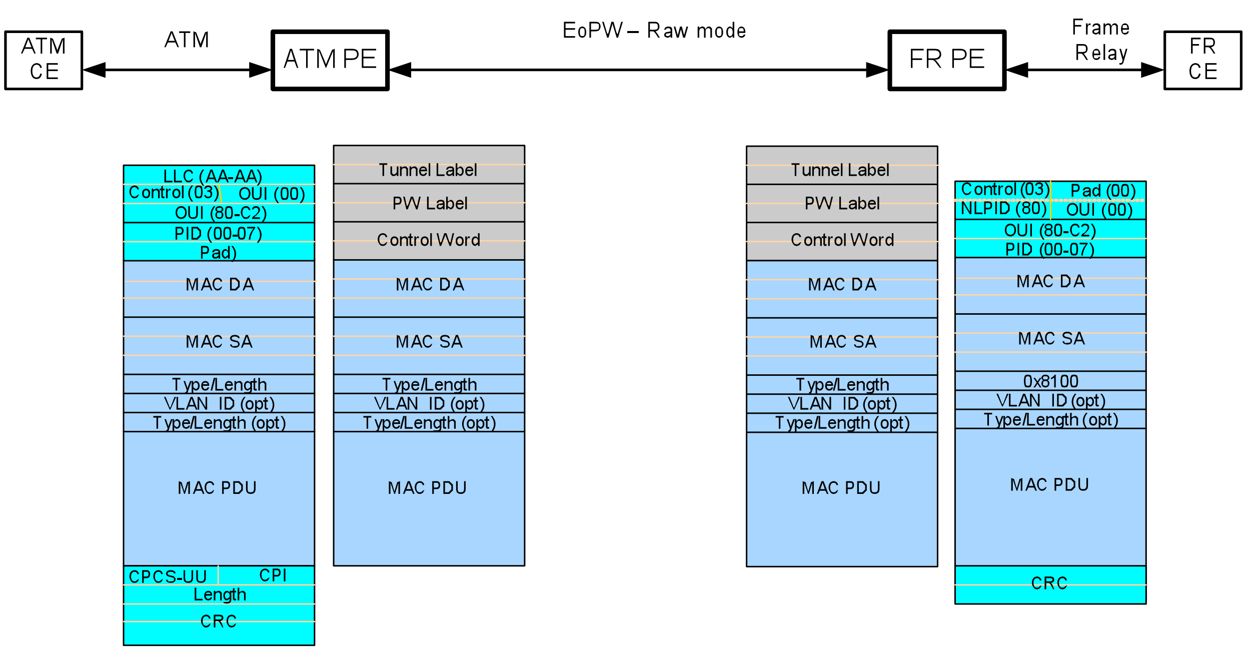

FIGURE 9. Frame Format for FR to ATM for Ethernet Encapsulation Interworking w/o FCS retention, Raw mode EoPW

6.4 PPP to Ethernet Bridge Interworking

Figure 10 shows the network scenarios for PPP to Ethernet interworking. EoMPLS on the pseudo wire is used for Ethernet service. In this case, Bridging Control Protocol [RFC 2878] can be used to provide Ethernet transport over PPP. The PE with the Ethernet link functions in the same way as a like-to-like scenario, and thus no additional modifications to the Ethernet PE are required. The pseudo wire here is also the same as the like-to-like scenario, that is, EoMPLS. The PE with the PPP link, in this case, terminates the PPP session completely, including both NCP and LCP layers. After extracting Ethernet frames from the EoMPLS pseudo wire, the PE encapsulates them using the BCP and sends them over the PPP session to the CE.

As described in section 5.1, only port-based mapping for PPP/HDLC VCs is required (e.g., all Ethernet traffic coming over the PPP/HDLC VC is associated with a single service instance, and is mapped into the Ethernet PW; no VLAN or VLAN-bundling mapping at the PPP AC is mandated).

The Ethernet PW can either be in raw-mode or tagged-mode, depending upon whether or not a service-delimiting tag needs to be carried over the EoPW (tagged mode must be used if a service-delimiting tag is carried). If tagged-mode is used, then the IWF will add a service-delimiting tag (S-VLAN) to the customer Ethernet frames, which can be either tagged or untagged. Regardless of whether the received customer frames are tagged or untagged, after addition of a S-VLAN tag to the customer frames by the PE, the resultant frame format and EtherType shall adhere to [IEEE 802.1ad].

Ethernet encapsulation over PPP is based upon [RFC-2878]. The BPDU frames are also carried in the same frame format. This format is the default format. These BPDU frames are identified based upon the MAC destination address in the frame format.

Note: If the PPP-PE is connected to an old BCP bridge and can support backwards compatibility, it must send BPDUs in the old format described in Appendix A of RFC 2878. Upon receiving a BPDU frame over the PPP link, the IWF at the PPP PE shall add the assigned BPDU MAC address as the MAC DA, and shall add a dummy unicast MAC address as the MAC SA before forwarding the BPDU frame to the egress PE over the EoPW. Support of this format is optional.

If the PW is of type tagged-mode, then upon receiving the Ethernet frames over the EoPW, the IWF in the Ethernet PE may or may not remove the service-delimiting tag, depending upon the type of the mapping at the Ethernet AC. If the mapping is of type port-based or VLAN-bundling, then the service-delimiting tags will be removed before the Ethernet frames are sent over the Ethernet AC. If the mapping is of type VLAN-mapping, then the service-delimiting tag (S-VLAN) will either get translated into the customer tag (C-VLAN) or it will get passed as is over the AC. If the AC is of type UNI, then the S-VLAN is translated to C-VLAN prior to frame transmission over the AC (frame format is translated from IEEE 802.1ad to IEEE 802.1q], but if the AC is of type NNI, then the S-VLAN tag is either transmitted over the AC as is or is mapped onto another S-VLAN tag (with IEEE 802.1ad frame format].

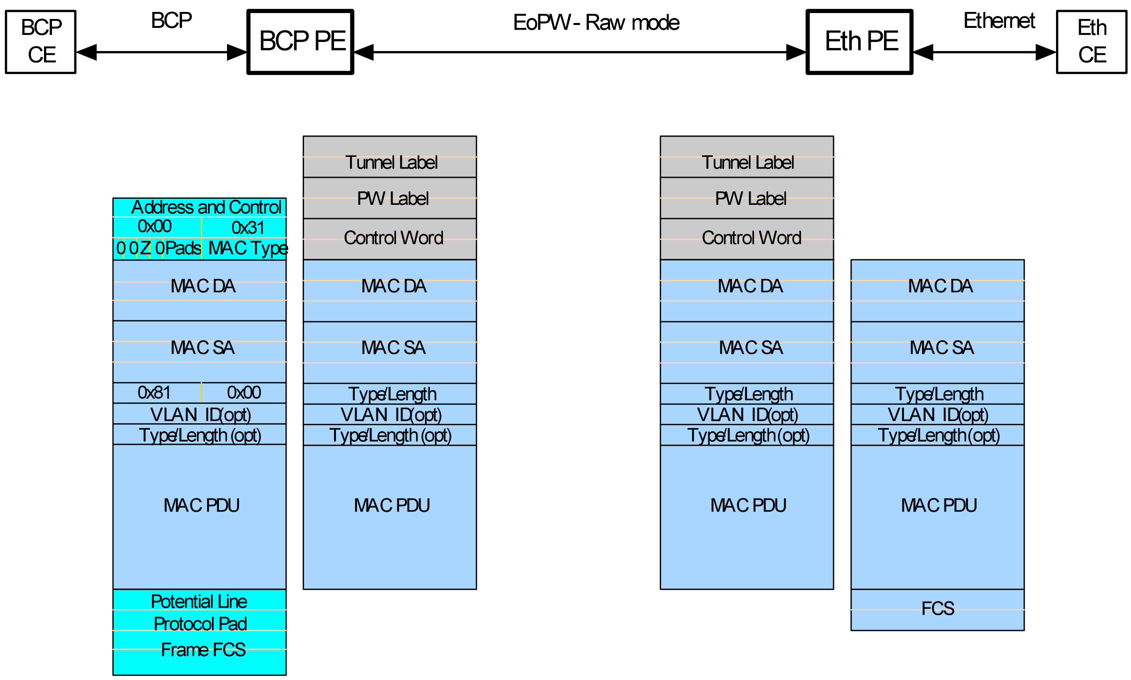

FIGURE 10. Frame format for PPP to Ethernet (Bridge) Interworking without FCS retention w Raw-mode EoPW

7 Traffic Management Interworking

In the context of this specification, Traffic Management (TM) interworking addresses the issue of different TM regimes applied at the Ethernet Service Instance (ESI) and at the attachment circuit (AC). By definition, this issue arises only when the AC type is not Ethernet. Appendix I addresses the case where the AC type is Frame Relay or ATM. The interworking functions and mappings in Appendix I are for information only.

8 PVC Management Interworking

Fault Management for Multiservice Interworking Version 1 [FM MS Inter] provides management-plane interworking for Ethernet services. There are two aspects to consider:

- When a PW fails: informing each CE device using the related OAM procedures for its corresponding AC
- When an AC fails: informing the remote PE device and subsequently the remote CE device regarding the failure of the local AC

The following subsections describe in detail each of the interworking scenarios described in section 6. In the case of the two-sided model, Ethernet frames are transported using Ethernet over MPLS pseudo wires. The following procedures of Fault Management for Multiservice specification version 1 are applicable:

8.1 Frame Relay - Ethernet Interworking

8.1.1 Attachment Circuit Defect Entry/Exit Procedures

See Sections 4.1.1.1; 4.1.2.1; 4.1.3.1 and 4.1.4.1 of Fault Management for Multiservice Specification version 1[FM-MSInt].

8.1.2 Pseudo Wire Entry/Exit Procedures

See Sections 4.2.1.1; 4.2.1.3; 4.2.2.1; 4.2.2.3; 4.2.3.1 and 4.2.4.1 of Fault Management for Multiservice Specification version 1 [FM-MSInt].

8.2 ATM - Ethernet Interworking

8.2.1 Attachment Circuit Defect Entry/Exit Procedures

See Sections 5.1 of Fault Management for Multiservice Specification version 1[FM-MSInt].

8.2.2 Pseudo Wire Entry/Exit Procedures

See Sections 5.2 of Fault Management for Multiservice Specification version 1 [FM-MSInt].

8.3 Frame Relay - ATM Interworking

8.3.1 PE attached to Frame Relay AC

See Section 8.1.

8.3.2 PE attached to ATM AC

See Section 8.2.

Annex A

Single-sided Model for Ethernet Interworking

(Normative)

This model results in several configurations of attachment circuits and pseudo-wire types. Those not addressed in the sections below are for further study.

A.1 ATM and Ethernet Attachment Circuits with ATM PW Encapsulation

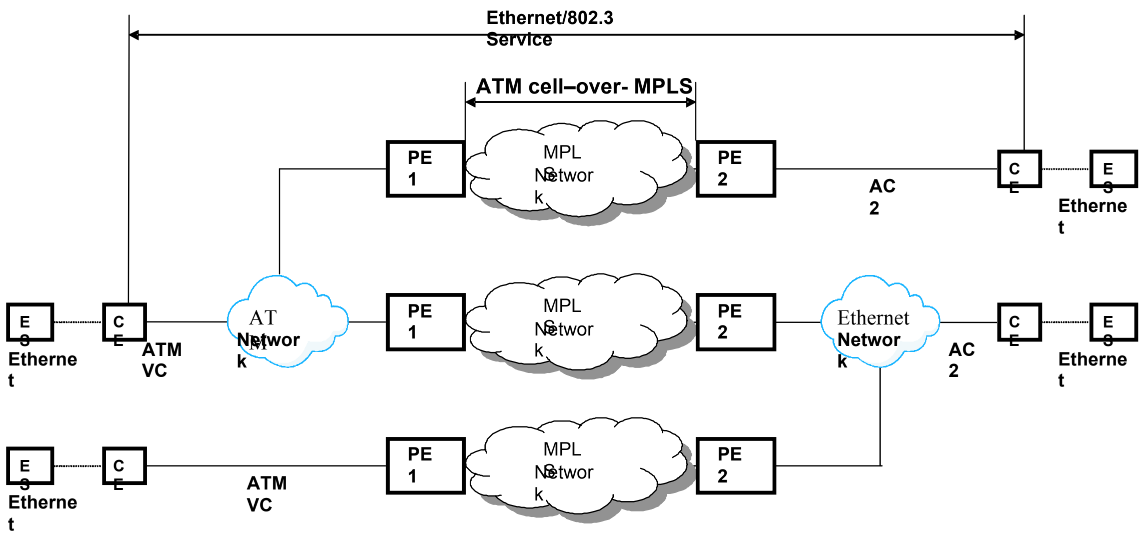

Figure 11: Single sided for Ethernet Interworking

This model should only be used when one of the Attachment Circuit (e.g., AC1) is of type ATM and furthermore, the corresponding PE (e.g., PE1) is not capable of performing Segmentation and Reassembly (SAR) functions to encapsulate and de-capsulate Ethernet frames. Alternatively, the network operator may want to extend the ATM AC to the remote PE for extending ATM OAM reach to the remote PE. Only under such circumstances is the use of this model advised.

In this case, PE1 transports the ATM cells over MPLS using a pseudo wire of type ATM cell mode. PE2 performs the SAR function on the ATM cells received/transmitted on the pseudo wire. It also performs Ethernet frame encapsulation or de-capsulation according to the procedures outlined in [RFC 2684].

In this model the Ethernet frames are first encapsulated using RFC 2684, then encapsulated using AAL5. The output ATM cells are again encapsulated inside the ATM-over-MPLS pseudo wire. This model creates additional overhead (see A.2).

As described in Section 5.1, only port-based mapping for ATM VCs is required (e.g., all Ethernet traffic coming over the ATM VC is associated with a single service instance and is mapped into the ATM PW).

Ethernet encapsulation over ATM VCs is based upon [RFC-2684]. It should be noted that the BPDU frames are treated differently at the Ethernet PE and have a different ATM encapsulation format (e.g., the PID is 0x0E and the encapsulation does not contain MAC addresses.). Therefore, upon receiving a BPDU frame over the ATM PW, the IWF at the Ethernet PE shall add a BPDU multicast MAC address as the MAC DA and shall add a dummy unicast MAC address as the MAC SA before forwarding the BPDU frame to the egress Ethernet AC. In the opposite direction, upon receiving a BPDU frame over the Ethernet AC, the IWF at the Ethernet PE shall remove the MAC header and shall encapsulate the frame based upon [RFC-2684] before delivering it over the ATM PW.

RFC 2684 supports the MAC FCS retention mode. This capability is used to support FCS retention. The non-FCS retention mode of operation is the default mode.

A.1.1 Frame Formatting and Delimiting

PE connected to the ATM CE:

It operates in the same way as in the case of ATM cell mode over MPLS for like-to-like service . It uses N-to-one encapsulation mode.

PE connected to the Ethernet CE:

- ATM to Ethernet direction: the Ethernet MAC frame is extracted from the ATM Bridged frame. The message delineation provided by AAL5 is used to identify frame boundaries. The ATM AAL5 SDU identifies the Ethernet MAC frame. Since MAC frames are included in the ATM AAL5 frames with bridged encapsulation, the ATM connection is terminated and the Ethernet MAC frame is delivered based upon the encapsulation in [RFC 2684].

- Ethernet to ATM direction: the Ethernet MAC frame is encapsulated into the ATM bridged encapsulation. frame format. The Ethernet MAC frame encapsulated using RFC 2684 is passed as AAL type 5 SDU. ITU-T Rec. I.363.5 is used for reassembly of the AAL5 PDU.

  The ATM VPI/VCI is configured to correspond to a VLAN-ID (in Ethernet VLAN-mode) or to all Ethernet frames (in Ethernet port-mode). The association is made at the time the PVC is provisioned.

  The congestion indication field (EFCI) and ATM User-to-user indication (AUU) bits are set to zero.

  Note: Mapping of the drop precedence representations between Ethernet and ATM is described in Appendix I, section 1.2.2.

The BPDU frames are treated differently at the ATM PE and have a different ATM encapsulation format (See section 6.2 for details).

A.1.2 Traffic Management and QoS

See Section 7.

A.1.3 PVC Management Interworking

Fault Management for Multiservice Interworking Version 1 [FM MS Inter] provides management-plane interworking for Ethernet services. Specifically, the procedures for ATM cell PWs apply.

Appendix I

Traffic Management Mapping

(Informative)

This Appendix describes: (i) mapping of the Ethernet traffic descriptor to FR or ATM, and vice versa; (ii) how discard eligibility representation can be mapped between Ethernet and Frame Relay or ATM. This Appendix addresses the case where the AC type is Frame Relay or ATM. The interworking functions and mappings in this Appendix are for information only.

Traffic Management functions, e.g. policing, marking, discarding, scheduling, and shaping typically operate on a given flow using parameters from a traffic descriptor model. Section I.1 describes traffic management interworking between Ethernet and Frame Relay. Section I.2 describes traffic management interworking between Ethernet and ATM.

The mappings described are generally inexact since the TM models of the various technologies differ. In some cases information is lost and in others approximations are necessary. A consequence of this is that conformance definitions applied at the ESI (Ethernet Service Instance) and at the AC may not find the same quantity of traffic conforming. It may be advisable in some cases, therefore, to exercise caution in the application of the mappings described in this appendix, and perhaps to include additional tolerance factors, as appropriate for the application. This appendix makes no specific recommendation with regard to any such factors.

Traffic Management mapping in this section is relevant to the case in which one of the attached circuits is of type ATM or FR

I.1 Ethernet and Frame Relay TM Interworking

This section describes Traffic Management interworking functionality to be performed at the PE in the directions from Ethernet to FR and from FR to Ethernet.

I.1.1 Mapping of Service Parameters

Ethernet traffic parameters [MEF10] and FR traffic parameters [I.370] are similar. The main difference between the two sets of parameters is that in Ethernet [MEF10] four parameters are defined, whereas in FR only three parameters are defined. The FR committed rate measurement interval Tc is computed.

The mapping of Ethernet and FR parameters are shown below:

Parameter Name	Ethernet Parameter	FR Parameter
Committed Information Rate	CIR_E (bits/sec)	CIR_F (bits/sec)
Committed Burst Size	CBS (octets)	B_c (bits)
Excess Information Rate	EIR (bits/sec)	not applicable
Excess Burst Size	EBS (octets)	B_e (bits)

I.1.1.1 Ethernet to FR Direction

In the Ethernet to FR direction the mapping of traffic descriptor parameters should follow the rules:

$$B_c = 8*CBS/a(L)$$

$$CIR_F = CIR_E/a(L)$$

$$B_e = 8*(EIR * (CBS/CIR_E))/a(L)$$

where a(L) is a correction factor that is a function of an Ethernet frame of length L. It is computed to account for the different amount of overhead introduced by Ethernet and FR (see Section I.1.3). The factors of 8 in the B_c and B_e equations account for the unit conversion from octets to bits.

In addition, the FR committed rate measurement interval, T_c, is defined [I.370] as (B_c/CIR_F), and so is also equal to ($8*CBS/CIR_E$).

Notice that B_e is defined not in terms of EBS, but rather in terms of the other three Ethernet traffic descriptor parameters. The implications of any mismatch between EBS and B_e should be considered, for example if an Ethernet traffic descriptor is used to configure a FR policer.

I.1.1.2 FR to Ethernet Direction

In the FR to Ethernet direction, the parameter mapping is straightforward and should follow the rules:

$$CBS = a(L)*B_c/8$$

$$EBS = a(L)*B_e/8$$

$$CIR_E = a(L)* CIR_F$$

$$EIR = a(L)*(B_e/T_c)$$

Note that there may be additional constraints that apply to the parameter mappings as well, e.g. CBS generally must satisfy a lower bound when CIR is non-zero (see Section 7.10.1 of [MEF 10]). Similar additional constraints apply between EBS and EIR.

I.1.2 Mapping of Discard Eligibility Indication

I.1.2.1 Ethernet to FR Direction

According to MEF10, Ethernet frames will be subject to a bandwidth profile at the UNI. Frames that are deemed in-profile will be marked as green while out-of-profile frames will be marked as yellow. The marking procedure itself is not defined in MEF10.

When color indication is available in the Ethernet frame (e.g. using user priority bits), at the interface between Ethernet and FR the following marking rules should be followed:

for green Ethernet frames, set DE = 0 in the FR frame header;

for yellow Ethernet frames, set DE = 1 in the FR frame header.

If the Ethernet frame color is not indicated in the Ethernet frame header, then the DE bit should always be set to 0 in the FR frame header.

I.1.2.2 FR to Ethernet Direction

If color indication is supported for Ethernet frames, then mapping discard eligibility in the FR to Ethernet direction is based upon the FR DE bit according to the following rules:

for DE = 0, set the Ethernet frame color to green;

for DE = 1, set the Ethernet frame color to yellow.

If color indication is not supported for Ethernet frames then no action is required for the mapping of the DE bit.

I.1.3 Overhead Calculation

As shown in the left side of Figure 7, a FR frame encapsulating an Ethernet frame, for transmission on a FR AC, adds 12 bytes of FR overhead (assuming a 2-byte Q.922 FR address). For a given Ethernet frame of length L bytes (following the usual convention of excluding the preamble and SFD), the correction factor a(L) used in Section I.1.1 is thus

$$a(L) = \frac{L}{L+12}.$$

Depending upon the use for which the mapping is employed, the value of L used in the correction factor might vary. For example, if the goal is a long-term approximate mapping, one might use a measure of the average Ethernet frame length, L_{ave}. If the goal is to ensure that the FR AC is provisioned with enough capacity to support all combinations of conforming Ethernet traffic, one might utilize a smaller, more conservative value, between L_{ave} and L_{min}. Similarly, there may be reasons to utilize a larger value, between L_{ave} and L_{max}.

I.2 Ethernet and ATM TM Interworking

This section describes Traffic Management interworking functionality to be performed at the PE in the directions from Ethernet to ATM and from ATM to Ethernet

I.2.1 Mapping of Service Parameters

I.2.1.1 Ethernet to ATM Direction

MEF10 does not define classes of service equivalent to ATM service categories [TM4.1]. However, the Ethernet and ATM traffic descriptors do share some general characteristics, e.g. each rate is defined in terms of a leaky bucket algorithm, with an associated tolerance parameter. The biggest difference between the approaches is that in Ethernet the highest allowable arrival rate is the sum of CIR and EIR, while in ATM that rate is captured in a single parameter, the Peak Cell Rate (PCR).

MEF10 states that performance parameters in terms of frame loss, delay, and delay variation are applicable only to Ethernet frames that are deemed in profile (green frames). No performance assurances are extended to out-of-profile Ethernet frames (yellow frames). ATM also distinguishes between cells that are provided performance (QoS) assurance and those that are not. Depending upon the specific conformance definition applied to an ATM virtual connection, QoS guarantees may be extended to different combinations of conforming CLP=0 and CLP=1 cells, or to no cells (in the case of UBR).

In the Ethernet to ATM direction the parameter mapping should use the following rules:

PCR = (CIR + EIR)/a(L)

SCR = CIR/a(L)

MCR = CIR/a(L)

MBS = 1 + {(8*CBS) / [(1 – CIR/LR)*a(L)]}

MFS = $\lceil$(maximum Ethernet frame size in octets + 16)/48$\rceil$, where $\lceil x \rceil$ is defined as the smallest integer that is greater than or equal to x.

In all cases, results need to be quantized to an integer number of cells or cells/second. The quantization direction depends upon whether the goal of the mapping is to be conservative, or to minimize the error, etc.

The correction factor a(L) accounts for the overhead in encapsulating an Ethernet frame into an AAL5 PDU and then segmenting the PDU into cells. It also converts the Ethernet parameters from units of bits/sec to the ATM parameter units of cells/sec. The factor of 8 in the MBS equation follows from the fact that CBS is defined in octets, not bits. See Section I.2.3 for the definition of a(L).

Every ATM virtual connection includes the PCR parameter. The other parameters above may or may not be relevant for a given connection, depending upon the ATM service category (they are never ALL employed on a single connection). See Table 2-1 in [TM 4.1] for a list of traffic descriptor parameters associated with each service category.

One ATM parameter that does not have a direct analog in Ethernet is the Cell Delay Variation Tolerance (CDVT), which is the leaky bucket tolerance associated with PCR. Guidelines for setting CDVT for an ATM AC are network-specific. The lack of a general mapping to CDVT reflects the inexact nature of the Ethernet-to-ATM traffic descriptor mapping, as noted in the introduction to this appendix.

The expression for MBS takes into account the fact that MBS is defined as the maximum burst size (cells) at PCR, while CBS is defined as the token bucket size in the Ethernet bandwidth profile. During a burst of Ethernet frames, tokens are consumed at rate LR, the effective Ethernet link rate on the AC, and are also replenished at rate CIR. LR is less than the physical line rate of the AC; it includes only the transmission bandwidth available for bits of Ethernet frames, and excludes bandwidth utilized for Layer 1 and Layer 2 (ATM) overhead. The ATM overhead is accounted in the a(L) correction term, as with the other parameters.

Note that there may be additional constraints that apply to the parameter mappings as well, e.g. the MBS for a GFR connection must satisfy a lower bound (see Section 4.5.5.2 of [TM 4.1]).

I.2.1.2 ATM to Ethernet Direction

The mapping of the ATM traffic descriptor to the Ethernet traffic descriptor depends upon the ATM service category. In the ATM to Ethernet direction the parameter mapping should use the following rules:

If the ATM service category is Constant Bit Rate:

CIR = a(L)*PCR,

CBS = a(L)*(CDVT*PCR)/8,

EIR = EBS = 0.

If the ATM service category is Variable Bit Rate (either real-time or non-real-time):

CIR = a(L)*SCR

CBS = a(L)*(MBS-1)*(1 – CIR/LR)/8

EIR = a(L)*(PCR-SCR)

EBS = a(L) * (CDVT * (PCR-SCR))/8.

If the ATM service category is Guaranteed Frame Rate:

CIR = a(L)*MCR

CBS = a(L)*(MBS-1)*(1 – CIR/LR)/8

EIR = a(L)*(PCR-MCR)

EBS = a(L) * (CDVT * (PCR-MCR))/8.

Maximum Ethernet frame size in octets = (MFS*48) - 16

If the ATM service category is Unspecified Bit Rate:

EIR = a(L)*PCR,

EBS = a(L)*(CDVT*PCR)/8,

CIR = CBS = 0.

If the ATM service category is Available Bit Rate (ABR), the mapping between ATM and Ethernet traffic descriptors is loose. ABR commits to a dynamic allowed cell rate, ACR(t) that varies between MCR and PCR over time. All traffic sent at or below ACR(t) is committed and it is never permissible to send excess traffic. A very conservative mapping would assign CIR and CBS as in the case of Guaranteed Frame Rate above, while setting EIR = EBS = 0, but that expressly foregoes the opportunity to take advantage of higher allowed rates over time. A mapping that takes advantage of a dynamic allowed rate would reassign CIR = a(L)*ACR(t) as ACR(t) varies, but a dynamic CIR is outside the scope of MEF10.

The term LR above is defined as in Section I.2.1.1.

Note that there may be additional constraints that apply to the parameter mappings as well, e.g. CBS generally must satisfy a lower bound when CIR is non-zero (see Section 7.10.1 of [MEF 10]). Similar additional constraints apply between EBS and EIR.

I.2.2 Mapping of Drop Precedence Indication:

I.2.2.1 Ethernet to ATM Direction

When color indication is available in the Ethernet frame (e.g. using user priority bits), at the interface between Ethernet and ATM the following marking rules should be followed

for green (in-profile) Ethernet frames, set CLP = 0 in all the ATM cells of the frame;

for yellow (out-of-profile) Ethernet frames, set CLP = 1 in all ATM cells of the frame.

Notes:

If the ATM service category is GFR or UBR, it is network-specific whether yellow frames must be sent in CLP=1 ATM cells (see Appendix VI.4 of [TM4.1] for a discussion of GFR implementation options). However, GFR does require that all cells of a given frame have the same CLP value.

If the ATM service category is ABR, it is not permissible to send an Ethernet frame in ATM cells with CLP = 1. Regardless of the color of the Ethernet frame, it is the responsibility of the ABR source node to transmit cells within the dynamic allowed cell rate, and all such cells must have CLP=0.

If Ethernet frame color is not indicated in the Ethernet frame header, then the CLP bit should be set to 0 in all ATM cells.

I.2.2.2 ATM to Ethernet Direction:

If any cell has CLP = 1, then the corresponding Ethernet frame will be marked yellow. If all cells of a given Ethernet frame have CLP = 0, the Ethernet frame is marked green. However, as noted above, it is network-specific whether the appearance of only CLP=0 cells within a GFR frame can be relied upon as an indication that the frame is within the connection's MCR.

If color indication is not supported for Ethernet frames, then no action is required for the mapping of the CLP bits.

I.2.3 Overhead Calculation

The overhead calculation in the case of ATM is more complex than that in FR. It includes three factors: 1) a factor to convert from bits (or bits/second) to cells (or cells/second); 2) a factor to count the overhead introduced by encapsulating the Ethernet frame into an AAL5 PDU; and 3) a factor to account for cell overhead.

The first factor is 424, i.e. there are 424 bits in a cell.

As shown in the left side of Figure 8, the encapsulation of the Ethernet frame into an AAL5 PDU adds between 16 and 63 octets (the pad is between 0 and 47 octets, and is chosen so that the AAL5 PDU length is an integer multiple of 48 octets). This variability is accounted for by using the "roundup" operator, $\lceil x \rceil$, defined as the smallest integer that is greater than or equal to x.

The third factor is 53/48, i.e. each 53 byte cell has 5 bytes of cell header and 48 bytes of AAL5 PDU payload.

Combining these factors, the correction term a(L) for an Ethernet frame of length L octets, is:

$$a(L) = (424) * (L / \lceil (L+16)/48 \rceil * 48) * (48/53)$$

$$= (L*8) / \lceil (L+16)/48 \rceil \quad \text{(bits/cell)}.$$

As noted in Section I.1.3, the value of L used in the correction factor will in general depend upon the use for which the mapping is employed.

END OF DOCUMENT