Soft Permanent Virtual Circuit Interworking between MPLS Pseudowires and ATM

MFAF.18.0.0

MFA Forum Technical Committee
September 2007

Editor:

George Swallow

Cisco Systems

For more information contact:

The MFA Forum
Suite 307
39355 California Street
Fremont, CA 94538 USA

Phone: +1 (510) 608-3997
FAX: +1 (510) 608-5917
E-Mail: info@mplsforum.org
WWW: http://www.mplsforum.org/

Revision History

Version	Change	Date
MFAF.x.0	Initial version	March 2007

Table of Contents

1 INTRODUCTION 1

2 TERMINOLOGY 1

3 SERVICE MEDIATION ARCHITECTURE 3

4 BACKGROUND ON IDENTIFIERS 4

4.1 PSEUDOWIRE IDENTIFIERS 4
4.2 ATM SPVC IDENTIFIERS 5

5 PROPOSED SOLUTION 6

5.1 MPLS / ATM INTERFACE 6
5.2 SIGNALING 7
5.3 MAPPING IDENTIFIERS 7
5.4 REASSEMBLY AND PAYLOAD IDENTIFICATION 11
5.5 CONFIGURATION AND PROCEDURES 11

6 REFERENCES 13

1 Introduction

In current Asynchronous Transfer Mode (ATM) deployments, Soft Permanent Virtual Connections (SPVCs) are used to provision both ATM Permanent Virtual Channel Connections (PVCC) and Permanent Virtual Path Connections (PVPC) and Frame Relay (FR) PVCCs.

Pseudowires over Multiprotocol Label Switching (MPLS) Packet Switched Networks (PSNs) are currently being introduced as a backbone technology to carry these same services. Mechanisms are being developed to enable a flexible provisioning model that incorporates elements of the SPVC model, namely that configuration of the end service exists only at the end-points. These mechanisms are described in the Pseudowire Control Protocol [1] and Provisioning Models and Endpoint Identifiers in L2VPN Signaling [4]. While Pseudowires are defined over PSNs other than MPLS, the scope of this document is limited to the case were the PSN is MPLS.

As services are migrated from ATM to PSNs, any reasonable deployment scenario mandates that there be a period of time (possibly protracted or permanent) in which services will need to be established and maintained between end-users where one end of a circuit is attached to an ATM network and the other end is attached to a PSN.

To facilitate the migration of ATM and Frame Relay SPVCs to a PSN carrying pseudowires, a means of interoperating ATM and LDP signaling needs to be defined. Further this interoperation must preserve the essential reasons for using SPVCs, namely, keeping configuration limited to the edge nodes supporting a particular connection and allowing the network to be able to recover in the event of the failure of any facility between those two edge nodes.

It may also useful to perform reassembly of AAL5 frames, particularly in the case of Frame Relay connections at the boundary between the ATM network and the MPLS PSN. This serves to reduce dataplane protocol overhead in the PSN.

Finally, any solution must not preclude any existing services. In particular, Frame Relay to ATM interworking is recognized to be in wide use.

This document describes the architecture and procedures for interworking Pseudowire Signaling and ATM Signaling across a UNI, AINI, or PNNI [2].

2 Terminology

AAL5 ATM Adaptation Layer 5

AC Access Circuit

AESA ATM End System Address

AFI Authority and Format Identifier

AGI Attachment Group Identifier

AI Attachment Identifier

AII Attachment Individual Identifier

AINI ATM Inter -Network Interface

ATM Asynchronous Transfer Mode

DLCI Data Link Connection Identifier

IE Information Element

FEC Forwarding Equivalence Class

FR Frame Relay

IANA Internet Assigned Numbers Authority

ICP Internet Code Point

IDI Initial Domain Identifier

IP Intern

L2E Layer 2 Edge

L2PE Layer 2 Provider Edge

LDP Label Distribution Protocol

MME MPLS Mediation Edge

MPE MPLS Provider Edge

MPLS Multi-Protocol Label Switching

NSAP Network Service Access Point

OSI Open Systems Interconnect

PNNI Private Network-Network Interface

PSN Packet Switched Network

PVCC Permanent Virtual Channel Connection

PVPC Permanent Virtual Path Connection

SAI Source Attachment Identifier

SAII Source Attachment Individual Identifier

SPVC Soft Permanent Virtual Connection

TAI Target Attachment Identifier

TAII Target Attachment Individual Identifier

TLV Type, Length, Value

UNI User Network Interface

VCI Virtual Connection Identifier

VPI Virtual Path Identifier

VPN Virtual Private Network

3 Service Mediation Architecture

Service mediation addresses network scenarios where a layer-2 connection is established between an edge device that is attached to a native layer-2 network and an MPLS edge device that is attached to an MPLS core network. The end-to-end layer-2 connection is built from two segments: one segment is in its native layer-2 form established using native layer-2 signaling and the other segment is an MPLS-based pseudowire using standards-based PWE3 signaling protocols.

We call such service a "**mediated**" service to indicate that the native layer-2 signaling is terminated at the MPLS device attached to the layer-2 network (which we refer as the MPLS Mediation Edge -MME). The MME mediates the service by mapping the native signaling to MPLS-based PWE3 signaling across the MPLS network.

Figure 1 shows the network reference diagram for Service Mediation. The service provider network consists of one or multiple native layer-2 networks attached to an MPLS network. Examples of layer-2 networks are Frame Relay and ATM networks. The MPLS network is composed of Provider Edge devices and internal Provider devices. For the purpose of service mediation, we partition the MPLS network Provider Edge devices into "MPLS Provider Edge" devices, denoted as MPE, and "MPLS Mediation Edge" devices, referred as MME. An MPE connects customer sites that are attached to the MPLS network, and an MME interconnects layer-2 networks to the MPLS network. Note that an MPE could also be an MME.

Similarly, we partition the layer-2 network devices into Layer-2 Provider Edge devices (L2PE) and Layer-2 Edge devices (L2E). A L2PE is a layer-2 switch attaching customer sites and an L2E is a layer-2 device that is attached to an MME. Examples of the L2E-MME interface are PNNI, AINI, etc. Note that an L2E can also be an L2PE.

The native layer-2 network is modeled as a layer-2 VPN, which is assigned a globally unique identifier within the MPLS network (i.e., VPN-ID, AGI, etc). The MPLS network is modeled as a layer-2 host with one or multiple layer-2-based addresses.

In order for the layer-2 signaling to reach the set of MMEs, the MPLS network needs to be reachable from the layer-2 network. This is accomplished by provisioning layer-2 prefix addresses routable within the layer-2 network. NSAP summary addresses are configured at the L2E (when the L2E-MME interface is AINI or UNI) or MME (when the L2E-MME interface is PNNI). The layer-2 network (L2PE) has a view of which MME nodes it can reach. The prefix address could represent the entire MPLS domain, distinct network regions, or specific nodes (MPEs).

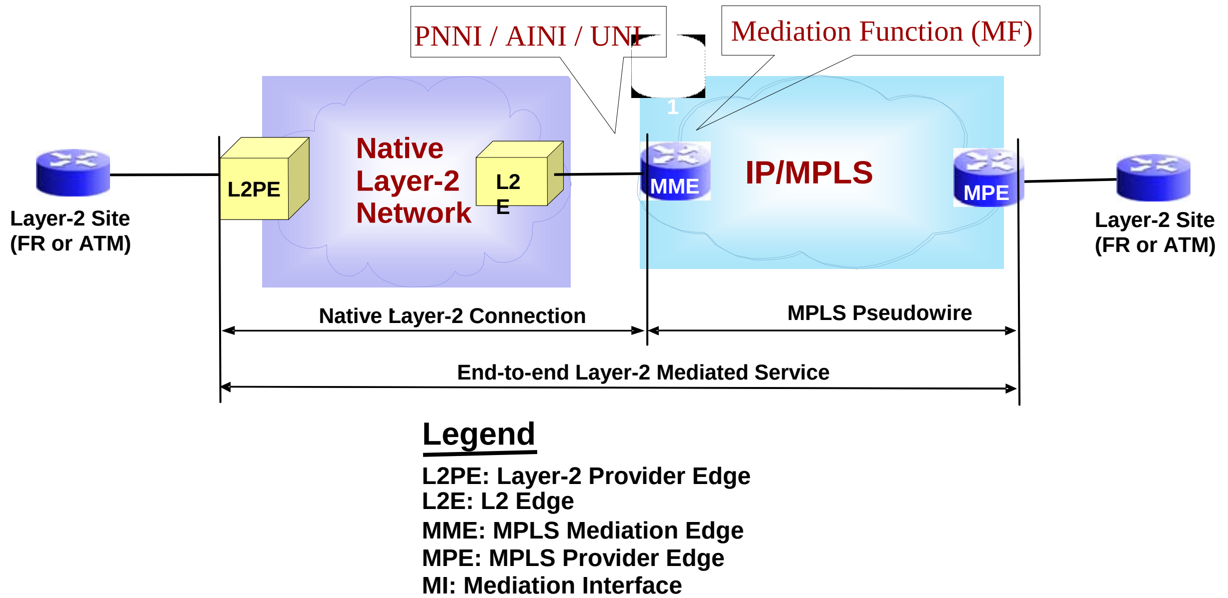


Figure 1: Service Mediation Network Reference

To setup a mediated service, the signaling protocols and the mediation function must be able to associate the Forwarder in L2PE with the Forwarder in the MPE. This is accomplished by encoding information that exactly identifies the target forwarder and the destination MPE in the MPLS network in the call setup messages originating from the layer-2 network. Within the ATM signaling, forwarder located on the MPE side will be identified using ATM related information (such as port number, VPI/VCI, or DLCI values) and the called layer-2 address must contain the IP address of the destination MPE.

This document defines support for ATM (and FR over ATM where either FRF.8.1 [11] or FRF.8.2 [12] may be used) as the layer-2 network, where the mediated connections are initiated only from the ATM network. Further it requires the ATM edge device to encode the remote MPLS provider edge IP address within the NSAP destination address.

4 Background on identifiers

4.1 Pseudowire Identifiers

Pseudowires serve to connect a pair of attachment circuits (ACs). In the context of this document these ACs are either Frame Relay DLCIs or ATM VPI/VCIs. On the ingress side, the AC is a VPI or VCI within the L2E-MME interface. The AC on the egress side is identified by an Attachment Identifier (AI) and the IP address of the egress MPE. AIs are defined in [1]. An AI is a logical reference for both the physical/logical port as well as the virtual circuit. That is, an AC is fully identified by a node-ID (IP address) and an AI.

The AI has further structure to designate groups of identifiers and individual identifiers within a group; these are called attachment group identifiers (AGI) and attachment individual identifiers (AII), respectively. When pairs of AIs are used in signaling, they are further designated by their role in the call, with the operative terms being source and target of the call. Thus we also have the terminology, source attachment identifier (SAI), source attachment individual identifier (SAII), target attachment identifier (TAI), and target attachment individual identifier (TAII). The source and target designations are only relevant from the perspective of the pseudowire control protocol. For example, a node receiving an LDP label mapping message from a remote node will swap the SAI and TAI values when it sends a label mapping message in the reverse direction back to the originating node.

Attachment Identifiers (AIs) are carried in the Generalized ID FEC Element of LDP. Each AI is encoded as two fields, the Attachment Group Identifier (AGI) and an Attachment Individual Identifier (AII), with each encoding using a type-length-value (TLV) format as defined in Section 5.2.2 of [1]. In particular:

> Note that the interpretation of a particular field as AGI, SAII, or TAII depends on the order of its occurrence. The type field identifies the type of the AGI, SAII, or TAII.

The rules for constructing the AGI and AII are left to the specification of applications and/or models.

4.2 ATM SPVC Identifiers

In ATM signaling [2], the identifying information of the attachment circuit at the destination interface consists of an ATM End-System Address (AESA) and the DLCI or VPI/VCI. The AESA identifies both the destination node and port where the end-user is attached. The DLCI or VPI/VCI are signaled in the Called party SPVC IE and are carried as literal values. The Called party SPVC IE has two formats depending on whether the service being signaled is a Frame Relay SPVC or an ATM SPVC. Furthermore, ATM SPVCCs and ATM SPVPCs are distinguished through the bearer class codepoint in the Broadband bearer capability IE.

AESAs are based on the NSAP address format. Figure 2 shows the generic format.

AFI	Authority and Format Identifier
IDI	Initial Domain Identifier
HO-DSP	High Order Domain Specific Part
ESI	End System Identifier
SEL	Selector

Figure 2: Generic AESA Format

Although many formats are permitted within AESAs, all ATM Forum defined formats include a six byte End System Identifier or ESI. The ESI's role is to identify a host. Typically, the first 13 bytes of the AESA are common for all end systems attached to an ATM node. This is the default behavior in PNNI. In this case, the ESI is used to differentiate between end systems attached to the same switch. From the point of view of the egress ATM switch, the ESI maps to a physical or logical port.

Thus common practice is to use the ESI to carry the port information.

5 Proposed Solution

5.1 MPLS / ATM Interface

The interface between the ATM network and the MPLS network can be any of the following:

- ATM Forum User Network Interface (UNI 4.1) [9]
- Private Network-Network Interface (PNNI) [2]
- ATM Inter-Network Interface (AINI) [10]

In the case of the UNI, there must be extensions to support the Called party soft PVPC or PVCC IE, if not all of UNI 4.1. (In this document we refer to the Called party soft PVPC or PVCC IE as simply the SPVC IE.) There may be extensions to support:

- the Calling party soft PVPC or PVCC IE,
- signaled VPs, using the "transparent VP service" codepoint for the bearer class in the Broadband bearer capability IE,
- crankback indication by setting the cause location in the Cause IE to a value other than "user", and

- frame discard indication using either the Frame Discard bits or the ATM adaptation layer parameters IE.

5.2 Signaling

In ATM, soft PVCs are statically defined across the UNI interfaces, but are signaled across the ATM network using PNNI signaling. For the signaling part, one edge node is configured to be active for the SPVC and the other is defined to be passive. That is, one end always initiates the call. ATM signaling creates a bidirectional circuit in a single pass. Bandwidth parameters for each direction of the circuit are carried in the setup message.

A paradigm for pseudowires that is analogous to the active/passive roles in SPVC setup above is known as single sided provisioning. These procedures are defined in Provisioning Models and Endpoint Identifiers in L2VPN Signaling [4], section 3.1.2. A small difference exists in this specification in that the "discovery" process occurs when an ATM setup message arrives.

It should be noted that pseudowire setup results in a pair of unidirectional labels assigned by two essentially independent label requests. The reverse label request is triggered by the forward label request. Furthermore, pseudowire signaling does not carry bandwidth parameters.

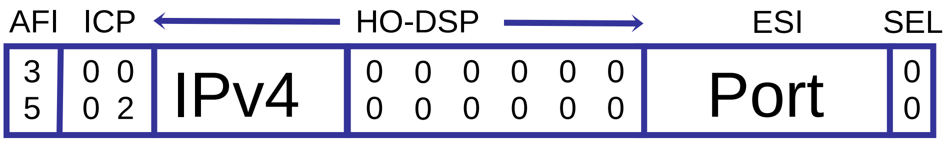
5.3 Mapping Identifiers

In Pseudowire Control Plane Signaling, an IP address identifies the egress node, and the (T)AI identifies the egress port and DLCI or VPI/VCI. In ATM, an AESA identifies both the egress node and port, and the DLCI or VPI/VCI is carried as a literal in the SPVC IE.

Two issues must be addressed. First a mapping between ATM and IP addresses is needed. Second, there must be a means of translating between the ATM port and SPVC IE and the pseudowire AIs.

5.3.1 Mapping Addresses

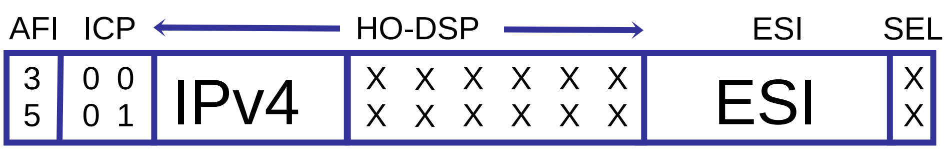
OSI Network Service Access Point Addresses include an Authority and Format Identifier (AFI). The AFI value 35 has been assigned to IANA. According to RFC4548 [6], when the AFI is set to 35 the two-octet Initial Domain Identifier (IDI) in Figure 2 encodes the Internet Code Point (ICP). Two ICP values are defined in [6], one each for IPv4 addresses and IPv6 addresses.. The codepoint 0x0001 indicates that an IPv4 address is encoded in the next 4 octets. The remaining octets are available to be assigned as the owner of the encoded IP address desires. For our purposes, the format shown in Figure 3 is recommended for use and required for implementation. In Figures 3, 4 and 5 each column represents one octet, with the upper digit being the hexadecimal value of the upper nibble and the lower digit representing the lower nibble.



AFI	Address Format Indicator
ICP	Internet Code Point
HO-DSP	High Order Domain Specific Part
ESI	End System Identifier
SEL	Selector

Figure 3: Recommended Use of NSAP with Embedded IPv4 Address

While it is common practice to carry the port number in the ESI field, we note that there are six unused bytes in the HO-DSP as well as the Selector Byte which could be used in a situation where the ESI is not available. In Figure 4, the available bytes (other than the ESI) are marked with Xs.



AFI	Authority and Format Identifier
ICP	Internet Code Point
HO-DSP	High Order Domain Specific Part
ESI	End System Identifier
SEL	Selector

Figure 4: Available bytes in an NSAP with Embedded IPv4 Address

RFC4548 also assigns ICP 0x0000 to indicate that an IPv6 address is encoded in the next 16 octets. In this format the only unused space is the Selector Byte. This allows for the identification of 256 ports. If more ports are needed, multiple addresses must be assigned.

AFI	Address Format Identifier
ICP	Internet Code Point
SEL	Selector
PT	Port

Figure 5: NSAP with Embedded IPv6 Address

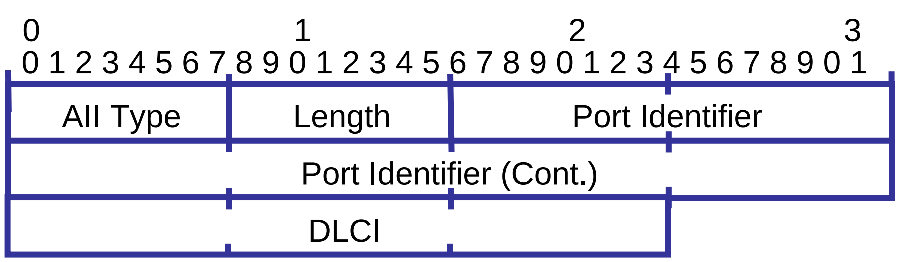
While it is expected that for IPv4 the ESI will commonly be used and for IPv6 the Selector Byte must be used, the discussion below simply refers to a generic port field.

5.3.2 Mapping the Port Identifier and SPVC IE to an AI

The AI is composed of an AGI and an AII. In the single-sided signaling procedures documented in [4], the AGI serves a specific purpose. Therefore, no special use of the AGI is made by this specification. Rather, the MME is configured with an AGI, and the procedures of [4] are followed.

The port identifier and SPVC IE values are reformatted to form the AII. The SPVC IE can carry either a Frame Relay DLCI or an ATM VPCI and optionally an ATM VCI. These are reformatted to one of two AII Types, one for Frame Relay and one for ATM. In the ATM format the presence or absence of a VCI is inferred from the length field.

5.3.2.1 Frame Relay AII Format



AII Type

0x05

Length

8 or 9 depending on the number of bytes in the DLCI field.

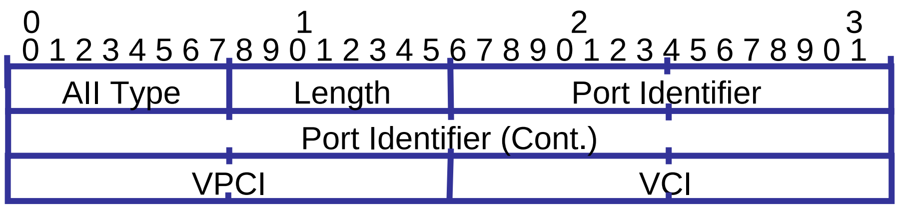
Port Identifier

From the destination NSAP address. By default, it is the ESI value. By configuration, the port identifier can be any combination of NSAP fields.

DLCI

For a 10-bit DLCI, the value of the DLCI as a 16-bit number, i.e. the value of the DLCI in the SPVC IE right justified and padded on the left with zeroes. For a 23-bit DLCI the value of the DLCI as a 24-bit number, i.e. the value of the DLCI in the SPVC IE right justified and padded on the left with a zero.

5.3.2.2 ATM AII Format



AII Type

0x06

Length

8 or 10 depending on the presence of the VCI field.

Port Identifier

From the destination NSAP address. By default it is the ESI value. By configuration, the port identifier can be any combination of NSAP fields.

VPCI

Value of the VPCI as a 16-bit number.

VCI

Value of the VCI as a 16-bit number. For SPVPCs this field is omitted.

5.4 Reassembly and Payload Identification

Reassembly at the MME is the recommended default behavior for Frame Relay and AAL5 services. In this case the PW encapsulation is recommended to be AAL5 SDU frame mode, as defined in Encapsulation Methods for Transport of Asynchronous Transfer Mode (ATM) Over MPLS Networks [5]. On a per ATM/PSN interface basis, an MME may be configured to not perform reassembly for Frame Relay. When reassembly does not occur at the MME, ATM SPVCCs and SPVPCs are encapsulated using one of the cell-mode encapsulations. The recommended encapsulation is ATM N-to-1 Cell Mode.

The encapsulation of the PW payload is signaled by indicating a PW Type as specified in [1]. When an MME is not capable of performing reassembly it is recommended that it be configured with a single PW Type to use for all pseudowires.

When an MME is capable of reassembly and is also capable of determining from the ATM signaling whether or not to perform reassembly, it is recommended that it be configured with a single PW Type for cell transport and a single PW Type for frame transport.

In some situations it is not possible for an MME that is capable of reassembly to determine if reassembly should be performed for a particular connection. In this case, it is recommended that the MME use the Wildcard PW Type as specified in [8]. This permits the MME to learn the PW Type from information configured at the MPE.

5.5 Configuration and Procedures

L2PEs: For each Permanent Virtual Connection, the L2PE is configured with the target AESA and the DLCI or VPI/VCI.

When the connection between the ATM and MPLS networks is an ANNI or UNI the L2Es are configured with the AESA prefix representing the set of MPEs reachable through its link(s) to MMEs. In the case of PNNI, this configuration occurs on the MME. Multiple prefixes may be configured to allow choices of optimum nodes to reach and to allow reachability to a larger set of nodes, should some other L2E or MME fail. The advertisement into the ATM network's routing protocol (e.g.PNNI) should be withdrawn if the associated link(s) have failed.

MPEs: An MPE is configured with the (T)AII of each of its ACs. Associated with each AC is a primary PW Type. Secondary PW Types may also be configured.

MMEs: An MME must be able to encode and parse the [to be assigned] AII types. Associated with each AII type is an AII format (used to form a TAII) and rules for how to extract the the IP address and port from the ATM called party address.

Configuration of PW Types for MMEs is described in section 5.4 above.

5.5.1 Procedures within the ATM Network

In an SPVC, one end is designated as the 'owner' and is responsible for initiating the connection. In order to simplify the interworking, this specification proposes that SPVCs always be initiated from the ATM side. This obviates any need to communicate bandwidth information across the PSN to the ATM network.

The L2E, as the owner of the connection, initiates PNNI signaling as it normally would. Finding a longest match associated with one or more of the AEs, it performs normal PNNI routing selection and sends a SETUP message which includes the SPVC IE.

When the SETUP message arrives at the L2E, it performs normal PNNI signaling processing and forwards the message across the PNNI, AINI or UNI to the MME.

5.5.2 Processing a SETUP message at the MME

When the MME receives a SETUP message, it performs ATM admission control. Additionally, the MME may perform additional checks to determine if it has the necessary resources to support the pseudowire connection in the forward direction. For example, in some network deployments it may determine if a PSN tunnel can be established in order to satisfy QoS or restoration requirements.

In the event that the call cannot be admitted, the MME SHOULD set an appropriate cause code, assuming that it is capable of supporting crankback procedures.

When the MME has successfully performed ATM admission control, it splices the call to a pseudowire using the signaling procedures of Provisioning Models and Endpoint Identifiers in L2VPN Signaling [4]. First, it extracts the destination IP address from the ATM called party address. It then determines if a LDP session exists with this node. If not, one is initiated. It then examines the SPVC IE to determine the type of service that is being requested. Based on the service type it selects AII type and format. It then extracts the port, VPI, VCI, and/or DLCI information, as appropriate to the service, and formats an AII. It formats an SAI using the same encoding rules, using the port the SETUP message was received on and the Connection Identifier. This AI becomes the handle that will be used to locate the context for this call. If an MME is capable of determining from the ATM signaling messages what PW Type to use it MAY insert that PW Type into the Label Mapping message. The means for making such a determination are beyond the scope of this document. Otherwise it simply uses whatever PW Type it is configured to use. It then inserts the SAI and TAI, and sends the Label Mapping message to the target node.

5.5.3 Procedures for the MPE

The MPE follows the procedures of [4] and [8] with no changes. That is, when the MPE receives a Label Mapping message, it uses the TAI to identify the interface and the DLCI or VPI/VCI. No decoding of the TAII is necessary; the AI and AC are simply configured as in [4].

If the forward label mapping completed successfully, it checks the PW Type. If the PW Type is "Wildcard" it responds with a Label Mapping message in the reverse direction indicating the configured primary PW Type. If the PW Type is not "Wildcard" and can be supported it responds with a Label Mapping message in the reverse direction indicating the received PW Type. Otherwise it MUST respond with a Label Release message with an LDP status code of "Generic Misconfiguration Error".

5.5.4 Call Completion at the MME

When the MME receives a Label mapping message from the target node, it uses the TAI (i.e. what it sent as the SAI) to locate the call context. It checks the PW Type and other call parameters to ensure that they can be supported. If they cannot be supported it MUST send a Label Release message, send a Label Withdrawal message and clear the ATM call indicating cause No. 63, "service or option not available, unspecified". Otherwise it completes the ATM call by sending a CONNECT message back to the L2PE.

6 References

[1] Martini, L., et al. *Pseudowire Setup and Maintenance Using the Label Distribution Protocol (LDP)*, IETF, RFC 4447, April 2006.

[2] ATM Forum, af-pnni-0055.001, *Private Network-Network Interface Specification Version 1.1*, April 2002.

[3] ATM Forum, str-cs-ipaddr-01.01, *IP-Based Addressing Version 1.0*, October 2003.

[4] Rosen, E. et al. *Provisioning, Autodiscovery, and Signaling in L2VPNs*, draft-ietf-l2vpn-signaling-08.txt, May 2006. (Note: in RFC Editor Queue)

[5] Martini, et al., *Encapsulation Methods for Transport of Asyncrhonous Transfer Mode (ATM) Over MPLS Networks*, RFC 4717, December 2006.

[6] Gray, E., et al., *Internet Code Point (ICP) Assignments for NSAP Addresses*, RFC 4548, May 2006.

[7] Martini, L., *IANA Allocations for Pseudowire Edge to Edge Emulation (PWE3)*, BCP 116, RFC 4446, April 2006.

[8] Swallow, G., and L. Martini, *Wildcard Pseudowire Type*, RFC 4863, May 2007.

[9] ATM Forum, af-sig-0061.002, *ATM User Network Interface(UNI) Signalling Specification Version 4.1*, April 2002.

[10] ATM Forum, af-cs-125.002, *ATM Inter-Network Interface (AINI) Specification Version 1.1*, September 2002.

[11] Frame Relay Forum, FRF.8.2, *Frame Relay/ATM PVC Service Interworking Implementation Agreement*, February 2000.

[12] Frame Relay Forum, FRF.8.2, *Frame Relay/ATM PVC Service Interworking Implementation Agreement*, February 2004.

END OF DOCUMENT