

Network Working Group
Request for Comments: 3916
Category: Informational

X. Xiao, Ed.
Riverstone Networks
D. McPherson, Ed.
Arbor Networks
P. Pate, Ed.
Overture Networks
September 2004

Requirements for Pseudo-Wire Emulation Edge-to-Edge (PWE3)

Status of this Memo

This memo provides information for the Internet community. It does not specify an Internet standard of any kind. Distribution of this memo is unlimited.

Copyright Notice

Copyright (C) The Internet Society (2004).

Abstract

This document describes base requirements for the Pseudo-Wire Emulation Edge to Edge Working Group (PWE3 WG). It provides guidelines for other working group documents that will define mechanisms for providing pseudo-wire emulation of Ethernet, ATM, and Frame Relay. Requirements for pseudo-wire emulation of TDM (i.e., "synchronous bit streams at rates defined by ITU G.702") are defined in another document. It should be noted that the PWE3 WG standardizes mechanisms that can be used to provide PWE3 services, but not the services themselves.

Table of Contents

1.	Introduction	2
1.1.	What Are Pseudo Wires?.	2
1.2.	Current Network Architecture.	3
1.3.	PWE3 as a Path to Convergence	4
1.4.	Suitable Applications for PWE3.	4
1.5.	Summary	4
2.	Terminology	5
3.	Reference Model of PWE3	6
4.	Packet Processing	7
4.1.	Encapsulation	7
4.2.	Frame Ordering.	8
4.3.	Frame Duplication	8
4.4.	Fragmentation	8

- 4.5. Consideration of Per-PSN Packet Overhead. 9
- 5. Maintenance of Emulated Services. 9
 - 5.1. Setup and Teardown of Pseudo-Wires. 9
 - 5.2. Handling Maintenance Message of the Native Services . . 10
 - 5.3. PE-initiated Maintenance Messages 10
- 6. Management of Emulated Services 12
 - 6.1. MIBs. 12
 - 6.2. General MIB Requirements. 12
 - 6.3. Configuration and Provisioning. 13
 - 6.4. Performance Monitoring. 13
 - 6.5. Fault Management and Notifications. 13
 - 6.6. Pseudo-Wire Connection Verification and Traceroute. . . 13
- 7. Faithfulness of Emulated Services 13
 - 7.1. Characteristics of an Emulated Service. 14
 - 7.2. Service Quality of Emulated Services. 14
- 8. Non-Requirements. 14
- 9. Quality of Service (QoS) Considerations 15
- 10. Inter-domain Issues 16
- 11. Security Considerations 16
- 12. Acknowledgments 17
- 13. References. 17
 - 13.1. Normative References. 17
 - 13.2. Informative References. 17
- 14. Authors' Addresses. 18
- 15. Full Copyright Statement. 19

1. Introduction

1.1. What Are Pseudo Wires?

Pseudo Wire Emulation Edge-to-Edge (PWE3) is a mechanism that emulates the essential attributes of a service such as ATM, Frame Relay or Ethernet over a Packet Switched Network (PSN). The required functions of PWs include encapsulating service-specific PDUs arriving at an ingress port, and carrying them across a path or tunnel, managing their timing and order, and any other operations required to emulate the behavior and characteristics of the service as faithfully as possible.

From the customer perspective, the PW is perceived as an unshared link or circuit of the chosen service. However, there may be deficiencies that impede some applications from being carried on a PW. These limitations should be fully described in the appropriate service-specific documents and Applicability Statements.

1.2. Current Network Architecture

The following sections give some background on where networks are today and why they are changing. It also talks about the motivation to provide converged networks while continuing to support existing services. Finally, it discusses how PWs can be a solution for this dilemma.

1.2.1. Multiple Networks

For any given service provider delivering multiple services, the current infrastructure usually consists of parallel or "overlay" networks. Each of these networks implements a specific service, such as Frame Relay, Internet access, etc. This is expensive, both in terms of capital expense and operational costs. Furthermore, the presence of multiple networks complicates planning. Service providers wind up asking themselves these questions:

- Which of my networks do I build out?
- How many fibers do I need for each network?
- How do I efficiently manage multiple networks?

A converged network helps service providers answer these questions in a consistent and economical fashion.

1.2.2. Transition to a Packet-Optimized Converged Network

In order to maximize return on their assets and minimize their operating costs, service providers often look to consolidate the delivery of multiple service types onto a single networking technology.

As packet traffic takes up a larger and larger portion of the available network bandwidth, it becomes increasingly useful to optimize public networks for the Internet Protocol. However, many service providers are confronting several obstacles in engineering packet-optimized networks. Although Internet traffic is the fastest growing traffic segment, it does not generate the highest revenue per bit. For example, Frame Relay traffic currently generates higher revenue per bit than native IP services do. Private line TDM services still generate even more revenue per bit than does Frame Relay. In addition, there is a tremendous amount of legacy equipment deployed within public networks that does not communicate using the Internet Protocol. Service providers continue to utilize non-IP equipment to deploy a variety of services, and see a need to interconnect this legacy equipment over their IP-optimized core networks.

1.3. PWE3 as a Path to Convergence

How do service providers realize the capital and operational benefits of a new packet-based infrastructure, while leveraging the existing equipment and also protecting the large revenue stream associated with this equipment? How do they move from mature Frame Relay or ATM networks, while still being able to provide these lucrative services?

One possibility is the emulation of circuits or services via PWs. Circuit emulation over ATM and interworking of Frame Relay and ATM have already been standardized. Emulation allows existing services to be carried across the new infrastructure, and thus enables the interworking of disparate networks.

Implemented correctly, PWE3 can provide a means for supporting today's services over a new network.

1.4. Suitable Applications for PWE3

What makes an application suitable (or not) for PWE3 emulation? When considering PWs as a means of providing an application, the following questions must be considered:

- Is the application sufficiently deployed to warrant emulation?
- Is there interest on the part of service providers in providing an emulation for the given application?
- Is there interest on the part of equipment manufacturers in providing products for the emulation of a given application?
- Are the complexities and limitations of providing an emulation worth the savings in capital and operational expenses?

If the answer to all four questions is "yes", then the application is likely to be a good candidate for PWE3. Otherwise, there may not be sufficient overlap between the customers, service providers, equipment manufacturers and technology to warrant providing such an emulation.

1.5. Summary

To maximize the return on their assets and minimize their operational costs, many service providers are looking to consolidate the delivery of multiple service offerings and traffic types onto a single IP-optimized network.

In order to create this next-generation converged network, standard methods must be developed to emulate existing telecommunications

formats such as Ethernet, Frame Relay, and ATM over IP-optimized core networks. This document describes requirements for accomplishing this goal.

2. Terminology

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALLNOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119.

Some terms used throughout this document are listed below.

Attachment Circuit (AC)

The physical or virtual circuit attaching a CE to a PE. An AC can be a Frame Relay DLCI, an ATM VPI/VCI, an Ethernet port, a VLAN, a HDLC link, a PPP connection on a physical interface, a PPP session from an L2TP tunnel, an MPLS LSP, etc.

Customer Edge (CE)

A device where one end of a service originates and/or terminates. The CE is not aware that it is using an emulated service rather than a native service.

Packet Switched Network (PSN)

Within the context of PWE3, this is a network using IP or MPLS as the mechanism for packet forwarding.

Provider Edge (PE)

A device that provides PWE3 to a CE.

Pseudo Wire (PW)

A mechanism that carries the essential elements of an emulated circuit from one PE to another PE over a PSN.

Pseudo Wire Emulation Edge to Edge (PWE3)

A mechanism that emulates the essential attributes of a service (such as a T1 leased line or Frame Relay) over a PSN.

Pseudo Wire PDU

A Protocol Data Unit (PDU) sent on the PW that contains all of the data and control information necessary to emulate the desired service.

PSN Tunnel

A tunnel across a PSN inside which one or more PWs can be carried.

3. Reference Model of PWE3

A pseudo-wire (PW) is a connection between two provider edge (PE) devices which connects two attachment circuits (ACs). An AC can be a Frame Relay DLCI, an ATM VPI/VCI, an Ethernet port, a VLAN, a HDLC link, a PPP connection on a physical interface, a PPP session from an L2TP tunnel, an MPLS LSP, etc.

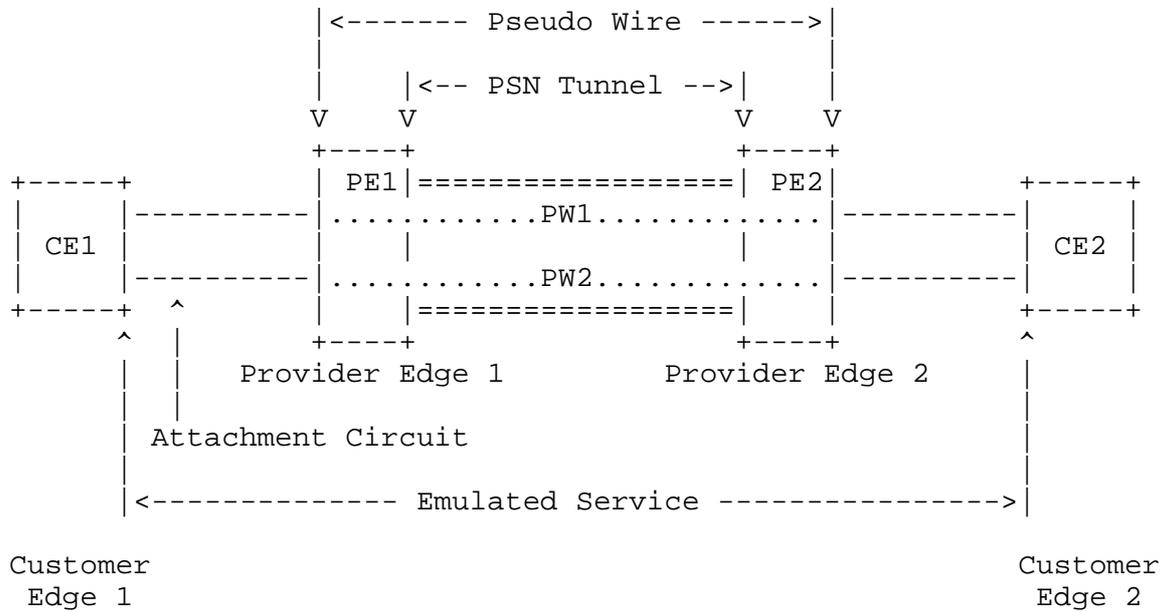


Figure 1: PWE3 Reference Model

During the setup of a PW, the two PEs will be configured or will automatically exchange information about the service to be emulated so that later they know how to process packets coming from the other end. After a PW is set up between two PEs, frames received by one PE from an AC are encapsulated and sent over the PW to the remote PE, where native frames are re-constructed and forwarded to the other CE. For a detailed PWE3 architecture overview, readers should refer to the PWE3 architecture document [PWE3_ARCH].

This document does not assume that a particular type of PWs (e.g., [L2TPv3] sessions or [MPLS] LSPs) or PSNs (e.g., IP or MPLS) is used. Instead, it describes generic requirements that apply to all PWs and PSNs, for all services including Ethernet, ATM, and Frame Relay, etc.

4. Packet Processing

This section describes data plane requirements for PWE3.

4.1. Encapsulation

Every PE MUST provide an encapsulation mechanism for PDUs from an AC. It should be noted that the PDUs to be encapsulated may or may not contain L2 header information. This is service specific. Every PWE3 service MUST specify what the PDU is.

A PW header consists of all the header fields in a PW PDU that are used by the PW egress to determine how to process the PDU. The PSN tunnel header is not considered as part of the PW header.

Specific requirements on PDU encapsulation are listed below.

4.1.1. Conveyance of Necessary L2 Header Information

The egress of a PW needs some information, e.g., which native service the PW PDUs belong to, and possibly some L2 header information, in order to know how to process the PDUs received. A PWE3 encapsulation approach MUST provide some mechanism for conveying such information from the PW ingress to the egress. It should be noted that not all such information must be carried in the PW header of the PW PDUs. Some information (e.g., service type of a PW) can be stored as state information at the egress during PW setup.

4.1.2. Support of Variable Length PDUs

A PWE3 approach MUST accommodate variable length PDUs, if variable length PDUs are allowed by the native service. For example, a PWE3 approach for Frame Relay MUST accommodate variable length frames.

4.1.3. Support of Multiplexing and Demultiplexing

If a service in its native form is capable of grouping multiple circuits into a "trunk", e.g., multiple ATM VCCs in a VPC or multiple Ethernet 802.1Q interfaces in a port, some mechanism SHOULD be provided so that a single PW can be used to connect two end-trunks. From encapsulation perspective, sufficient information MUST be carried so that the egress of the PW can demultiplex individual circuits from the PW.

4.1.4. Validation of PW-PDU

Most L2 frames have a checksum field to assure frame integrity. Every PWE3 service MUST specify whether the frame's checksum should be preserved across the PW, or should be removed at the ingress PE and then be re-calculated and inserted at the egress PE. For protocols such as ATM and FR, the checksum covers link-local information such as the circuit identifiers (e.g., FR DLCI or ATM VPI/VCI). Therefore, such checksum MUST be removed at the ingress PE and recalculated at the egress PE.

4.1.5. Conveyance of Payload Type Information

Under some circumstances, it is desirable to be able to distinguish PW traffic from other types of traffic such as IPv4 or IPv6 or OAM. For example, if Equal Cost Multi-Path (ECMP) is employed in a PSN, this additional distinguishability can be used to reduce the chance that PW packets get misordered by the load balancing mechanism. Some mechanism SHOULD provide this distinguishability if needed. Such mechanism MAY be defined in the PWE3 WG or other WGs.

4.2. Frame Ordering

When packets carrying the PW PDUs traverse a PW, they may arrive at the egress out of order. For some services, the frames (either control frames only or both control and data frames) must be delivered in order. For such services, some mechanism MUST be provided for ensuring in-order delivery. Providing a sequence number in the PW header for each packet is one possible approach to detect out-of-order frames. Mechanisms for re-ordering frames may be provided by Native Service Processing (NSP) [PWE3_ARCH] but are out of scope of PWE3.

4.3. Frame Duplication

In rare cases, packets traversing a PW may be duplicated. For some services, frame duplication is not allowed. For such services some mechanism MUST be provided to ensure that duplicated frames will not be delivered. The mechanism may or may not be the same as the mechanism used to ensure in-order frame delivery.

4.4. Fragmentation

If the combined size of the L2 payload and its associated PWE3 and PSN headers exceeds the PSN path MTU, the L2 payload may need to be fragmented (Alternatively the L2 frame may be dropped). For certain native service, fragmentation may also be needed to maintain a control frame's relative position to the data frames (e.g., an ATM PM

cell's relative position). In general, fragmentation has a performance impact. It is therefore desirable to avoid fragmentation if possible. However, for different services, the need for fragmentation can be different. When there is potential need for fragmentation, each service-specific PWE3 document MUST specify whether to fragment the frame in question or to drop it. If an emulated service chooses to drop the frame, the consequence MUST be specified in its applicability statement.

4.5. Consideration of Per-PSN Packet Overhead

When the L2 PDU size is small, in order to reduce PSN tunnel header overhead, multiple PDUs MAY be concatenated before a PSN tunnel header is added. Each encapsulated PDU still carries its own PW header so that the egress PE knows how to process it. However, the benefit of concatenating multiple PDUs for header efficiency should be weighed against the resulting increase in delay, jitter and the larger penalty incurred by packet loss.

5. Maintenance of Emulated Services

This section describes maintenance requirements for PWE3.

5.1. Setup and Teardown of Pseudo-Wires

A PW must be set up before an emulated circuit can be established, and must be torn down when an emulated circuit is no longer needed. Setup and teardown of a PW can be triggered by a command from the management plane of a PE, or by Setup/Teardown of an AC (e.g., an ATM SVC), or by an auto-discovery mechanism.

Every PWE3 approach MUST define some setup mechanism for establishing the PWs. During the setup process, the PEs need to exchange some information (e.g., to learn each other's capability). The setup mechanism MUST enable the PEs to exchange all necessary information. For example, both endpoints must agree on methods for encapsulating PDUs and handling frame ordering. Which signaling protocol to use and what information to exchange are service specific. Every PWE3 approach MUST specify them. Manual configuration of PWs can be considered as a special kind of signaling and is allowed.

If a native circuit is bi-directional, the corresponding emulated circuit can be signaled "Up" only when the associated PW and PSN tunnels in both directions are functional.

5.2. Handling Maintenance Message of the Native Services

Some native services have mechanisms for maintenance purpose, e.g., ATM OAM and FR LMI. Such maintenance messages can be in-band (i.e., mixed with data messages in the same AC) or out-of-band (i.e., sent in a dedicated control circuit). For such services, all in-band maintenance messages related to a circuit SHOULD be transported in-band just like data messages through the corresponding PW to the remote CE. In other words, no translation is needed at the PEs for in-band maintenance messages. In addition, it MAY be desirable to provide higher reliability for maintenance messages. The mechanisms for providing high reliability do not have to be defined in the PWE3 WG.

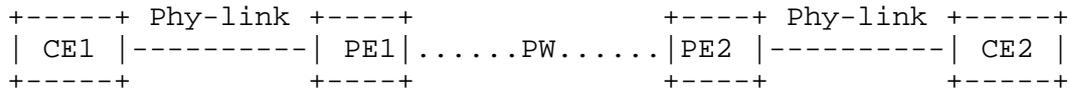
Out-of-band maintenance messages between a CE and a PE may relate to multiple ACs between the CE and the PE. They need to be processed at the local PE and possibly at the remote PE as well. If a native service has some out-of-band maintenance messages, the corresponding emulated service MUST specify how to process such messages at the PEs. In general, an out-of-band maintenance message is either translated into an in-band maintenance message of the native service or a PWE-specific maintenance message for every AC related to that out-of-band message. As an example, assume the ACs between a CE and a PE are some ATM VCCs inside a VPC. When a F4 AIS [UNI3.0] from the CE is received by the PE, the PE should translate that F4 AIS into a F5 AIS and send it to the remote CE for every VCC. Alternatively, the PE should generate a PWE-specific maintenance message (e.g., label withdrawal) to the remote PE for every VCC. When the remote PE receives such a PWE-specific maintenance message, it may need to generate a maintenance message of the native service and send it to the attached CE.

5.3. PE-initiated Maintenance Messages

A PE needs to initiate some maintenance messages under some circumstances without being triggered by any native maintenance messages from the CE. These circumstances are usually caused by fault, e.g., a PW failure in the PSN or a link failure between the CE and the PE.

The reason the PEs need to initiate some maintenance messages under a fault condition is because the existence of a PW between two CEs would otherwise reduce the CEs' maintenance capability. This is illustrated in the following example. If two CEs are directly connected by a physical wire, a native service (e.g., ATM) can use notifications from the lower layer (e.g., the physical link layer) to

assist its maintenance. For example, an ATM PVC can be signaled "Down" if the physical wire fails. However, consider the following scenario.



If the PW between PE1 and PE2 fails, CE1 and CE2 will not receive physical link failure notification. As a result, they cannot declare failure of the emulated circuit in a timely fashion, which will in turn affect higher layer applications. Therefore, when the PW fails, PE1 and PE2 need to initiate some maintenance messages to notify the client layer on CE1 and CE2 that use the PW as a server layer. (In this case, the client layer is the emulated service). Similarly, if the physical link between PE1-CE1 fails, PE1 needs to initiate some maintenance message(s) so that the client layer at CE2 will be notified. PE2 may need to be involved in this process.

In the rare case when a physical wire between two CEs incurs many bit errors, the physical link can be declared "Down" and the client layer at the CEs be notified. Similarly, a PW can incur packet loss, corruption, and out-of-order delivery. These can be considered as "generalized bit error". Upon detection of excessive "generalized bit error", a PW can be declared "Down" and the detecting PE needs to initiate a maintenance message so that the client layer at the CE is notified.

In general, every emulated service MUST specify:

- * Under what circumstances PE-initiated maintenance messages are needed,
- * Format of the maintenance messages, and
- * How to process the maintenance messages at the remote PE.

Some monitoring mechanisms are needed for detecting such circumstances, e.g., a PW failure. Such mechanisms can be defined in the PWE3 WG or elsewhere.

Status of a group of emulated circuits may be affected identically by a single network incidence. For example, when the physical link between a CE and a PE fails, all the emulated circuits that go through that link will fail. It is desirable that a single maintenance message be used to notify failure of the whole group of emulated circuits connected to the same remote PE. A PWE3 approach MAY provide some mechanism for notifying status changes of a group of emulated circuits. One possible approach is to associate each

emulated circuit with a group ID while setting up the PW for that emulated circuit. In a maintenance message, that group ID can be used to refer to all the emulated circuits in that group.

If a PE needs to generate and send a maintenance message to a CE, the PE MUST use a maintenance message of the native service. This is essential in keeping the emulated service transparent to the CEs.

The requirements stated in this section are aligned with the ITU-T maintenance philosophy for telecommunications networks [G805] (i.e., client layer/server layer concept).

6. Management of Emulated Services

Each PWE3 approach SHOULD provide some mechanisms for network operators to manage the emulated service. These mechanisms can be in the forms described below.

6.1. MIBs

SNMP MIBs [SMIV2] MUST be provided for managing each emulated circuit as well as pseudo-wire in general. These MIBs SHOULD be created with the following requirements.

6.2. General MIB Requirements

New MIBs MUST augment or extend where appropriate, existing tables as defined in other existing service-specific MIBs for existing services such as MPLS or L2TP. For example, the ifTable as defined in the Interface MIB [IFMIB] MUST be augmented to provide counts of out-of-order packets. A second example is the extension of the MPLS-TE-MIB [TEMIB] when emulating circuit services over MPLS. Rather than redefining the tunnelTable so that PWE can utilize MPLS tunnels, for example, entries in this table MUST instead be extended to add additional PWE-specific objects. A final example might be to extend the IP Tunnel MIB [IPTUNMIB] in such a way as to provide PWE3-specific semantics when tunnels other than MPLS are used as PSN transport. Doing so facilitates a natural extension of those objects defined in the existing MIBs in terms of management, as well as leveraging existing agent implementations.

An AC MUST appear as an interface in the ifTable.

6.3. Configuration and Provisioning

MIB Tables MUST be designed to facilitate configuration and provisioning of the AC.

The MIB(s) MUST facilitate intra-PSN configuration and monitoring of ACs.

6.4. Performance Monitoring

MIBs MUST collect statistics for performance and fault management.

MIBs MUST provide a description of how existing counters are used for PW emulation and SHOULD not replicate existing MIB counters.

6.5. Fault Management and Notifications

Notifications SHOULD be defined where appropriate to notify the network operators of any interesting situations, including faults detected in the AC.

Objects defined to augment existing protocol-specific notifications in order to add PWE functionality MUST explain how these notifications are to be emitted.

6.6. Pseudo-Wire Connection Verification and Traceroute

For network management purpose, a connection verification mechanism SHOULD be supported by PWs. Connection verification as well as other alarming mechanisms can alert network operators that a PW has lost its remote connection. It is sometimes desirable to know the exact functional path of a PW for troubleshooting purpose, thus a traceroute function capable of reporting the path taken by data packets over the PW SHOULD be provided.

7. Faithfulness of Emulated Services

An emulated service SHOULD be as similar to the native service as possible, but NOT REQUIRED to be identical. The applicability statement of a PWE3 service MUST report limitations of the emulated service.

Some basic requirements on faithfulness of an emulated service are described below.

7.1. Characteristics of an Emulated Service

From the perspective of a CE, an emulated circuit is characterized as an unshared link or circuit of the chosen service, although service quality of the emulated service may be different from that of a native one. Specifically, the following requirements MUST be met:

- 1) It MUST be possible to define type (e.g., Ethernet, which is inherited from the native service), speed (e.g., 100Mbps), and MTU size for an emulated circuit, if it is possible to do so for a native circuit.
- 2) If the two endpoints CE1 and CE2 of emulated circuit #1 are connected to PE1 and PE2, respectively, and CE3 and CE4 of emulated circuit #2 are also connected to PE1 and PE2, then the PWs of these two emulated circuits may share the same physical paths between PE1 and PE2. But from each CE's perspective, its emulated circuit MUST appear as unshared. For example, CE1/CE2 MUST NOT be aware of existence of emulated circuit #2 or CE3/CE4.
- 3) If an emulated circuit fails (either at one of the ACs or in the middle of the PW), both CEs MUST be notified in a timely manner, if they will be notified in the native service (see Section 5.3 for more information). The definition of "timeliness" is service-dependent.
- 4) If a routing protocol (e.g., IGP) adjacency can be established over a native circuit, it MUST be possible to be established over an emulated circuit as well.

7.2. Service Quality of Emulated Services

It is NOT REQUIRED that an emulated service provide the same service quality as the native service. The PWE3 WG only defines mechanisms for providing PW emulation, not the services themselves. What quality to provide for a specific emulated service is a matter between a service provider (SP) and its customers, and is outside scope of the PWE3 WG.

8. Non-Requirements

Some non-requirements are mentioned in various sections of this document. Those work items are outside scope of the PWE3 WG. They are summarized below:

- Service interworking;

In Service Interworking, the IWF (Interworking Function) between two dissimilar protocols (e.g., ATM & MPLS, Frame Relay & ATM, ATM & IP, ATM & L2TP, etc.) terminates the protocol used in one network and translates (i.e., maps) its Protocol Control Information (PCI) to the PCI of the protocol used in other network for User, Control and Management Plane functions to the extent possible.

- Selection of a particular type of PWs;
- To make the emulated services perfectly match their native services;
- Defining mechanisms for signaling the PSN tunnels;
- Defining how to perform traffic management on packets that carry PW PDUs;
- Providing any multicast service that is not native to the emulated medium.

To illustrate this point, Ethernet transmission to a multicast IEEE-48 address is considered in scope, while multicast services like [MARS] that are implemented on top of the medium are out of scope;

9. Quality of Service (QoS) Considerations

Some native services such as ATM can offer higher service quality than best effort Internet service. QoS is therefore essential for ensuring that emulated services are compatible (but not necessarily identical) to their native forms. It is up to network operators to decide how to provide QoS - They can choose to rely on over-provisioning and/or deploy some QoS mechanisms.

In order to take advantage of QoS mechanisms defined in other working groups, e.g., the traffic management schemes defined in DiffServ WG, it is desirable that some mechanisms exists for differentiating the packets resulted from PDU encapsulation. These mechanisms do not have to be defined in the PWE3 approaches themselves. For example, if the resulted packets are MPLS or IP packets, their EXP or DSCP field can be used for marking and differentiating. A PWE3 approach MAY provide guidelines for marking and differentiating.

The applicability of PWE3 to a particular service depends on the sensitivity of that service (or the CE implementation) to delay/jitter etc and the ability of the application layer to mask them. PWE3 may not be applicable to services that have severe constraints in this respect.

10. Inter-domain Issues

PWE is a matter between the PW end-points and is transparent to the network devices between the PW end-points. Therefore, inter-domain PWE is fundamentally similar to intra-domain PWE. As long as PW end-points use the same PWE approach, they can communicate effectively, regardless of whether they are in the same domain. Security may become more important in the inter-domain case and some security measure such as end-point authentication MAY be applied. QoS may become more difficult to deliver too, as one service provider has no control over another service provider's provisioning and traffic management policy. To solve the inter-domain QoS problem, service providers have to cooperate. Once they agree at a contractual level to provide high quality of service to certain traffic (e.g., PWE traffic), the mechanisms defined in other working groups, e.g., Diffserv WG, can be used.

Inter-domain PSN tunnels are generally more difficult to set up, tear down and maintain than intra-domain ones. But that is an issue for PSN tunneling protocols such as MPLS and L2TPv3 and is outside the scope of PWE3.

11. Security Considerations

The PW end-point, PW demultiplexing mechanism, and the payloads of the native service can all be vulnerable to attack. PWE3 should leverage security mechanisms provided by the PW Demultiplexer or PSN Layers. Such mechanisms SHOULD protect PW end-point and PW Demultiplexer mechanism from denial-of-service (DoS) attacks and spoofing of the native data units. Preventing unauthorized access to PW end-points and other network devices is generally effective against DoS attacks and spoofing, and can be part of protection mechanism. Protection mechanisms SHOULD also address the spoofing of tunneled PW data. The validation of traffic addressed to the PW Demultiplexer end-point is paramount in ensuring integrity of PW encapsulation. Security protocols such as IPsec [RFC2401] can be used.

12. Acknowledgments

The authors would like to acknowledge input from M. Aissaoui, M. Bocci, S. Bryant, R. Cohen, N. Harrison, G. Heron, T. Johnson, A. Malis, L. Martini, E. Rosen, J. Rutemiller, T. So, Y. Stein, and S. Vainshtein.

13. References

13.1. Normative References

- [IFMIB] McCloghrie, K. and F. Kastenholz, "The Interfaces Group MIB", RFC 2863, June 2000.
- [SMIV2] McCloghrie, K., Perkins, D., and J. Schoenwaelder, "Structure of Management Information Version 2 (SMIV2)", STD 58, RFC 2578, April 1999.

13.2. Informative References

- [G805] "Generic Functional Architecture of Transport Networks", ITU-T Recommendation G.805, 2000.
- [IPTUNMIB] Thaler, D., "IP Tunnel MIB", RFC 2667, August 1999.
- [L2TPv3] Lau, J., Townsley, M., and I. Goyret, et al., "Layer Two Tunneling Protocol (Version 3)", Work in Progress, June 2004.
- [MARS] Armitage, G., "Support for Multicast over UNI 3.0/3.1 based ATM Networks", RFC 2022, November 1996.
- [MPLS] Rosen, E., Viswanathan, A., and R. Callon, "Multiprotocol Label Switching Architecture", RFC 3031, January 2001.
- [PWE3_ARCH] S. Bryant and P. Pate, et. al., "PWE3 Architecture", Work in Progress, March 2004.
- [RFC2401] Kent, S. and R. Atkinson, "Security Architecture for the Internet Protocol", RFC 2401, November 1998.
- [TEMIB] Srinivasan, C., Viswanathan, A., and T. Nadeau, "Multiprotocol Label Switching (MPLS) Traffic Engineering (TE) Management Information Base (MIB)", RFC 3812, June 2004.
- [UNI3.0] ATM Forum, "ATM User-Network Interface Specification Version 3.0", Sept. 1993.

14. Authors' Addresses

XiPeng Xiao (Editor)
Riverstone Networks
5200 Great America Parkway
Santa Clara, CA 95054

EMail: xxiao@riverstonenet.com

Danny McPherson (Editor)
Arbor Networks

EMail: danny@arbor.net

Prayson Pate (Editor)
Overture Networks
507 Airport Boulevard, Suite 111
Morrisville, NC, USA 27560

EMail: prayson.pate@overturenetworks.com

Vijay Gill
AOL Time Warner

EMail: vijaygill9@aol.com

Kireeti Kompella
Juniper Networks, Inc.
1194 N. Mathilda Ave.
Sunnyvale, CA 94089

EMail: kireeti@juniper.net

Thomas D. Nadeau
Cisco Systems, Inc.
300 Beaver Brook Drive
Boxborough, MA 01719
EMail: tnadeau@cisco.com

Craig White
Level 3 Communications, LLC.
1025 Eldorado Blvd.
Broomfield, CO, 80021

EMail: Craig.White@Level3.com

15. Full Copyright Statement

Copyright (C) The Internet Society (2004).

This document is subject to the rights, licenses and restrictions contained in BCP 78, and except as set forth therein, the authors retain all their rights.

This document and the information contained herein are provided on an "AS IS" basis and THE CONTRIBUTOR, THE ORGANIZATION HE/S HE REPRESENTS OR IS SPONSORED BY (IF ANY), THE INTERNET SOCIETY AND THE INTERNET ENGINEERING TASK FORCE DISCLAIM ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Intellectual Property

The IETF takes no position regarding the validity or scope of any Intellectual Property Rights or other rights that might be claimed to pertain to the implementation or use of the technology described in this document or the extent to which any license under such rights might or might not be available; nor does it represent that it has made any independent effort to identify any such rights. Information on the IETF's procedures with respect to rights in IETF Documents can be found in BCP 78 and BCP 79.

Copies of IPR disclosures made to the IETF Secretariat and any assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of such proprietary rights by implementers or users of this specification can be obtained from the IETF on-line IPR repository at <http://www.ietf.org/ipr>.

The IETF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights that may cover technology that may be required to implement this standard. Please address the information to the IETF at ietf-ipr@ietf.org.

Acknowledgement

Funding for the RFC Editor function is currently provided by the Internet Society.

