Pseudo-Wire Edge-to-Edge (PWE3)
Internet Draft
Expires: July 2004

Thomas D. Nadeau
Monique Morrow
Cisco Systems, Inc

Peter Busschbach
Lucent Technologies
Editors

January 2004

Pseudo Wire (PW) OAM Message Mapping
draft-nadeau-pwe3-oam-msg-map-04.txt

Status of this Memo

Abstract

This document enumerates the OAM message mapping from pseudo wire emulated edge-to-edge services over MPLS and IP transport networks to their native attached services.

Table of Contents

1. Scope

This document covers the mapping of Pseudo Wire failures to error messages in the emulated services and the mapping of failures on Attachment Circuits (AC) to PW Status messages.

This document covers both PWE over MPLS PSN and PWE over IP PSN.

This document does not cover Service Interworking, i.e. the cases in which the native service at one side of the PW is different from the native service at the other side.

2. Terminology

AIS	Alarm Indication Signal
AOM	Administration, Operation and Maintenance
BDI	Backward Defect Indication
CC	Continuity Check
CE	Customer Edge
CPCS	Common Part Convergence Sublayer
DLC	Data Link Connection
FDI	Forward Defect Indication
FRBS	Frame Relay Bearer Service
IWF	Interworking Function

LB	Loopback
NE	Network Element
OAM	Operations and Maintenance
PE	Provider Edge
PW	Pseudowire
PSN	Packet Switched Network
RDI	Remote Defect Indicator
SDU	Service Data Unit
VCC	Virtual Channel Connection
VPC	Virtual Path Connection

The rest of this document will follow the following convention:

If LSP-Ping is run over a PW as described in [VCCV] it will be referred to as VCCV-Ping.

If BFD is run over a PW as described in [VCCV] it will be referred to as VCCV-BFD.

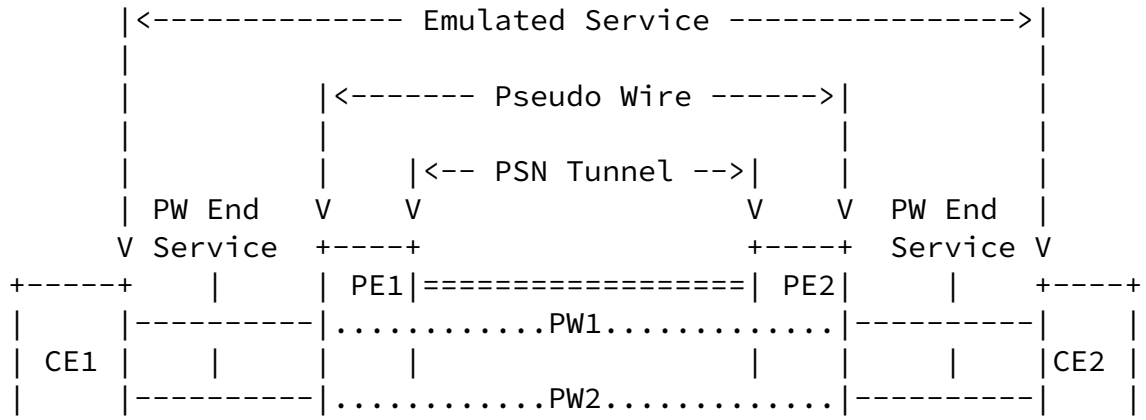
3. Introduction

This document describes how PW failures can be detected; how alarm information is exchanged between PEs; and how faults detected in pseudo-wires are mapped to OAM messages native to the emulated services and vice versa.

The objective of this document is to standardize the behavior of PEs with respects to failures on PWs and ACs, so that there is no ambiguity about the alarms generated and consequent actions undertaken by PEs in response to specific failure conditions.

3.1 Network Reference Model

Figure 1 illustrates the network reference model for point-to-point PWs.



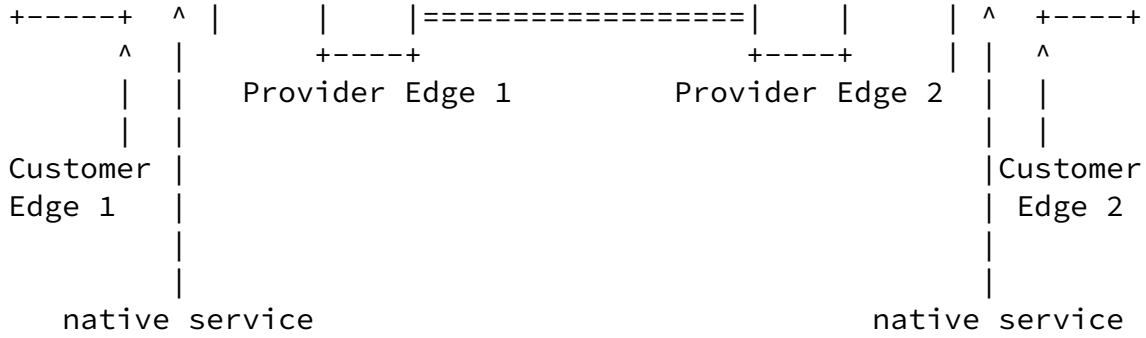


Figure 1: PWE3 Network Reference Model

This document discusses scenarios in which a single native service is emulated over a Pseudo Wire. Specifically, it discusses how the PEs translate PW failures (i.e. failures within PEs or between PEs) to error messages of the native service (i.e. between PE and CE)and vice

versa, in accordance with the guidelines set out in [OAM REQ].

4. PW Failures

This section describes possible PW failures, ways to detect them and consequent actions.

4.1 Failures

Possible failures that impact PWs are the following.

. Loss of connectivity between ingress and egress PE

. Control session failures between ingress and egress PE

In case of MPLS there are additional failures (see also [ITU-T Y.1710]). E.g.

. PW mislabeling, which could be due to a failure on the ingress PE or due to an over-writing of the PW label value somewhere along the way

. Label swapping or LSP mismerge in the PSN network, which could result in the termination of a PW at the wrong egress PE.

. Unintended self-replication (e.g. due to loops or denial-of-Service attacks)

4.1.1 Packet Loss

If PEs use sequence numbers for a specific Pseudo Wire, they have the

ability to detect packet loss. The question is at which point packet loss is so severe that a PW failure should be declared.

[CONGESTION] discusses possible mechanisms to detect congestion between PWs. The translation of packet loss to PW Failure should be discussed in the general framework of congestion control and is

therefore <TBD>.

4.2 Fault Detection

4.2.1 Available Fault Detection Tools

To detect the failures listed in 4.1, Service Providers have a variety of options available:

PSN Fault Detection Mechanisms:

For PWE3 over an IP PSN, with L2TP as encapsulation protocol, the fault detection mechanisms described in [L2TPv3] apply. Furthermore, the tools Ping and Traceroute, based on ICMP Echo Messages (see [ICMP] apply.

For PWE3 over an MPLS PSN, several tools can be used. E.g.: LSP-Ping and LSP-Traceroute (as defined in [LSPPING]) and Bi-directional Forwarding Detection ([BFD]). Furthermore, if RSVP-TE is used to setup the PSN Tunnels between ingress and egress PE, the hello protocol can be used to detect loss of connectivity (see [RSVP-TE]).

PW specific tools:

[VCCV] describes how LSP-Ping and BFD can be used over individual PWs. When used as such, we will refer to them as VCCV-Ping and VCCV-BFD respectively.

4.2.2 Tool Applicability

The discussion below is intended to give some perspective how tools mentioned in the previous section can be used to detect failures.

Observations:

. Tools like LSP-Ping and BFD can be run periodically or on demand. If used for failure detection (as opposed to diagnostic usage), they must be run periodically.

. Control protocol failures (e.g. detected through L2TPÆs Keep-alive messages or the Hello messages used in RSVP-TE) can be used to detect many network failures. However, control protocol failures do not necessarily coincide with data plane failures. Therefore, a

fault detection mechanism in the data plane is required to protect against all potential data plane failures.

. For PWE3 over an MPLS PSN, it may seem more effective to run a fault detection mechanism over a PSN Tunnel instead of over every individual PW within that PSN Tunnel. However in case the PSN traffic is distributed over Equal Cost Multi Paths (ECMP), it may be difficult to guarantee that PSN OAM messages follow the same path as a specific PW. A Service Provider might therefore decide to focus on fault detection over PWs.

. In MPLS networks, execution of LSP Ping would detect MPLS label errors, since it requests the receiving node to match the label with the original FEC that was used in the LSP set up. BFD can also be used since it relies on discriminators. A label error would result in a mismatch between the expected discriminator and the actual discriminator in the BFD control messages.

. For PWE3 over an MPLS PSN, PEs could detect PSN label errors through the execution of LSP-Ping. However, the same can be achieved with VCCV-Ping or VCCV-BFD. If, due to a label error in the PSN, a PW would be terminated on the wrong egress PE, PEs would detect this through the execution of VCCV-Ping and VCCV-BFD. In addition, VCCV-ping and VCCV-BFD would detect PW Label errors.

Based on these observations, it is clear that a Service Provider has the disposal of a variety of tools. There are many factors that influence which combination of tools best meets its needs.

4.3 Alarm Messages and Consequent Actions

When a PE detects a PW failure, it SHOULD inform its peer, by using:

. For PWE3 on MPLS PSN, PW Status messages as defined in [CONTROL].

. For PWE3 on IP PSN, L2TPv3 messages Stop Control-Connection Notification (SCCN) and Call Disconnect Notify (CDN) as defined in [L2TPv3]

Furthermore, in either case, if VCCV-BFD is used, the diagnostic code in the VCCV-BFD Control message can be used to exchange alarm information.

In general, PW Status messages or L2TPÆs SCCN and CDN should be used

to communicate failures. VCCV-BFD alarm indications should only be used in specific cases, as explained in 4.6.

Both PEs will translate the PW alarms to the appropriate failure

indications on the affected ACs. The exact procedures depend on the emulated protocols and will be discussed in the next sections.

4.4 The Use of PW Status

PW Status messages are used to report the following failures:

. Failures detected through fault detection mechanisms in the MPLS PSN

. Failures detected through VCCV (except for VCCV-BFD)

. Failures within the PE that result in an inability to forward traffic between ACs and PW

If the PW failure is related to one forwarding direction only, the PE shall either use "PW Receive Fault" or "PW Transmit Fault". In all other cases it shall use "PW Not Forwarding".

Besides reporting PW failures, PW status is used to propagate AC failures. When and how to use those messages is dependent on the emulated protocol and will be explained in the subsequent paragraphs (5.3 and 6.3).

4.5 The Use of L2TP SCCN and CDN

[L2TPv3] describes the use of SCCN and CDN messages to exchange alarm information between PEs. Like PW Status, SCCN and CDN messages shall be used to report the following failures:

. Failures detected through fault detection mechanisms in the IP PSN

. Failures detected through VCCV (except for VCCV-BFD)

. Failures within the PE that result in an inability to forward traffic between ACs and PW

In L2TP, the Set-Link-Info (SLI) message is used to convey failures on the ACs.

4.6 The Use of BFD Diagnostic Codes

[BFD] defines a set of diagnostic codes that partially overlap with failures that can be communicated through PW Status messages or L2TPÆs SCCN and CDN. To avoid ambiguous situations, these messages SHOULD be used for all failures that are detected through means other

than BFD.

For VCCV-BFD, therefore, only the following diagnostic codes apply:

0 û- No Diagnostic
1 û- Control Detection Time Expired
3 û- Neighbor Signaled Session Down
7 û- Administratively Down

[VCCV] states that, when used over PWs, the asynchronous mode of BFD should be used. Diagnostic code 2 (Echo Function Failed) does not apply to the asynchronous mode, but to the Demand Mode.

All other BFD diagnostic codes refer to failures that can be communicated through PW Status or L2TP SCCN and CDN.

The VCCV-BFD procedures are as follows:

When the downstream PE (PE1) does not receive control messages from the upstream PE (PE2) during a certain number of transmission intervals (a number provisioned by the operator), it declares that the PW in its receive direction is down. PE1 sends a message to PE2 with H=0 (i.e. "I do not hear you") and with diagnostic code 1. In turn, PE2 declares the PW is down in its transmit direction and it uses diagnostic code 3 in its control messages to PE2.

When a PW is taken administratively down, the PEs will exchange PW Status messages with code "Pseudo Wire Not Forwarding" or L2TP CDN messages with code "Session disconnected for administrative reasons". In addition, exchange of BFD control messages MUST be suspended. To that end, the PEs MUST send control messages with H=0 and diagnostic code 7.

Note: According to [BFD], control messages with an incorrect discriminator field must be discarded. However, since such an occurrence might be caused by swapped or mismerged LSPs, it would be better if a new diagnostic code were introduced to discriminate between missing control packets and packets with an incorrect discriminator.

4.7 PW Failure Entry and Exit Procedures

PWs can fail in a single direction or in both directions. PEs SHOULD keep track of the status of each individual direction. In other words, a PE SHOULD be able to distinguish between the following states: "PW UP", "PW Transmit Direction Down", "PW Receive Direction Down", "PW Receive and Transmit Down".

The next two sections discuss under which conditions a PE enters and exits these states. To avoid an unnecessarily complicated

description, only the states "PW UP" and "PW DOWN" are discussed without further analysis whether it applies to one or two directions of the PW.

4.7.1 PW Down

A PE will consider a PW down if one of the following occurs

. It detects a local failure

. It detects Loss of Connectivity or a Label Error on the PW

. It receives a message from its peer indicating a PW failure, which could be one of the following:

  o PW Status indicating "PW Receive Fault"; "PW Transmit

Fault"; or "PW not forwarding"

- o An L2TP SCCN or CDN message

- o A BFD Control message with diagnostic code "Neighbor Signaled Session Down"

Note that if the control session between the PEs fails, the PW is torn down and needs to be re-established. As a consequence, control session failure leads to disappearance of the PW, not to a PW Down state.

4.7.2 PW Up

When a PE determines that all previously existing failures have disappeared, it SHOULD send a message to its peer to indicate this. E.g. if the original failure was conveyed through a PW Status message, the PE should send a PW Status message indicating "PW Forwarding"

When a PE receives a PW Status message indicating "PW Forwarding", while it still considers a PW down, and if all previously existing failures, if any, have disappeared, it SHOULD respond with a PW Status message indicating "PW Forwarding".

A PE will exit the PW down state when the following conditions are true:

. All failures it had previously detected have disappeared

. It has received a PW Status message from its peer indicating "PW Forwarding".

[BFD] and [L2TPv3] define the procedures to exit the PW Down state if

the original failure notification was done through BFD or L2TP messages, respectively.

5. Frame Relay Encapsulation

5.1 Frame Relay Management

The management of Frame Relay Bearer Service (FRBS) connections can be accomplished through two distinct methodologies:

1. Based on ITU-T Q.933 Annex A, Link Integrity Verification procedure, where STATUS and STATUS ENQUIRY signaling messages are sent using DLCI=0 over a given UNI and NNI physical link. [ITU-T Q.933]

2. Based on FRBS LMI, and similar to ATM ILMI where LMI is common in private Frame Relay networks.

   In addition, ITU-T I.620 addresses Frame Relay loopback, but the deployment of this standard is relatively limited. [ITU-T I.620]

It is possible to use either, or both, of the above options to manage Frame Relay interfaces. This document will refer exclusively to Q.933 messages.

The status of any provisioned Frame Relay PVC may be updated through:

. STATUS messages in response to STATUS ENQUIRY messages, these are mandatory.

. Optional unsolicited STATUS updates independent of STATUS ENQUIRY (typically under the control of management system, these updates can be sent periodically (continuous monitoring) or only upon detection of specific defects based on configuration.

In Frame Relay, a DLC is either up or down. There is no distinction between different directions.

5.2 Mapping PW failures to Frame Relay OAM messages

In case a PE keeps track of the status of individual Frame Relay PVCs (which often, but not necessarily, coincides with the usage of 1-1 encapsulation mode), the following procedures apply:

When a PE determines the PW is down it will indicate this on the ACs through STATUS messages that indicate that the affected FR PVCs

linked to that PW are "inactive".

When the PE determines that the PW is up, it will indicate this on the AC through STATUS messages that indicate that the FR PVCs linked to that PW are "active".

In case of pure port mode, STATUS ENQUIRY and STATUS messages are transported transparently over the PW. A PW Failure will therefore result in timeouts at the Frame Relay devices at one or both sites of the emulated interface.

5.3 Frame Relay Attachment Circuit Failures

As explained in [CONTROL], if a PE detects that a Frame Relay PVC is "inactive", as defined in [ITU-T Q933] Annex A.5, it will convey this information to its peer. The remote PE SHOULD generate the corresponding errors and alarms on the egress Frame Relay PVC

For PWE3 over MPLS PSN, a PE that detects a failure on its AC shall send a PW Status message indicating both "AC Receive Fault" and "AC Transmit Fault".

For PWE3 over IP PSN, a PE that detects a failure on its AC shall send an L2TP Set-Link Info (LSI) message with a Circuit Status Attribute Value Pair (AVP) indicating "inactive".

6. ATM Encapsulation

6.1 ATM Management

ATM management and OAM mechanisms are much more evolved than those of Frame Relay. There are five broad management-related categories, including fault management (FT), Performance management (PM), configuration management (CM), Accounting management (AC), and Security management (SM). ITU-T Recommendation I.610 describes the functions for the operation and maintenance of the physical layer and the ATM layer, that is, management at the bit and cell levels ([ITU-T I.610]). Because of its scope, this document will concentrate on ATM fault management functions. Fault management functions include the following:

1) Alarm indication signal (AIS)
2) Remote Defect indication (RDI).
3) Continuity Check (CC).
4) Loopback (LB)

Some of the basic ATM fault management functions are described as follows: Alarm indication signal (AIS) sends a message in the same direction as that of the signal, to the effect that an error has been

detected.

Remote defect indication (RDI) sends a message to the transmitting terminal that an error has been detected. RDI is also referred to as the far-end reporting failure. Alarms related to the physical layer are indicated using path AIS/RDI. Virtual path AIS/RDI and virtual channel AIS/RDI are also generated for the ATM layer.

OAM cells (F4 and F5 cells) are used for the control of virtual paths and virtual channels with regard to their performance and availability. F4 cells are used to monitor a VPC, F5 cells for a VCC. OAM cells in the F4 and F5 flows are used for monitoring a segment of the network and end-to-end monitoring. OAM cells in F4 flows have the same VPI as that of the connection being monitored. OAM cells in F5 flows have the same VPI and VCI as that of the connection being monitored. The AIS and RDI messages of the F4 and F5 flows are sent to the other network nodes via the VPC or the VCC to which the message refers. The type of error and its location can be indicated in the OAM cells. Continuity check is another fault management function. To check whether a VCC that has been idle for a period of time is still functioning, the network elements can send continuity-check cells along that VCC.

6.2 Mapping PW Failures to ATM OAM

In case a PE keeps track of the status of individual ATM VPCs or VCCs, this behavior is specified in [PWEATM]:

In the VPC case, when a PW Failure is detected, the PE downstream from the failure MUST generate F4 AIS on the related ACs. In the VCC case, when a PW Failure is detected, the PE downstream from the failure SHOULD generate F5 AIS on the related ACs.

In case of transparent mapping, where the PE does not keep track of the status of individual ATM VPCs or VCCs, it is possible that a PE

does not know which VPCs and/or VCCs are active. In such a case there is a need for another fault indication mechanism on the AC. This is beyond the scope of this document.

6.3 ATM Attachment Circuit Failures

When an ingress PE detects a failure on an AC, it SHOULD generate F4 and/or F5 AIS on the PW towards the egress PE. The far end of the affected VC or VP segment will generate RDI, which is transported transparently over the PW in the reverse direction.

As stated in [CONTROL], there MAY exist implementations that do not transport OAM cells transparently.

For PWE3 over MPLS PSN, when an ingress PE detects a failure on an AC and it is not able to send OAM cells over the PW, it MUST send a PW Status message indicating "AC Receive Fault". On reception of this message, the egress PE MUST generate AIS on the related ATM VPCs or VCCs.

The far end of the affected VC or VP segment will generate RDI. When the egress PE receives an RDI signal over an AC and it is not able to transmit OAM cells, it MUST send a PW Status message indicating "AC Transmit Fault". On reception of this message, the ingress PE MUST generate RDI cells on the related VPCs and VCCs.

For PEW3 over IP PSN, when an ingress PE detects a failure on its AC and it is not able to send OAM cells over the PW, it MUST send an L2TP Set-Link Info (LSI) message with a Circuit Status AVP indicating "inactive". On reception of this message, the egress PE MUST generate AIS on the related ATM VPCs or VCCs. When the egress PE receives an RDI signal over an AC and it is not able to transmit OAM cells, it MUST send a L2TP Set-Link Info (LSI) message with a Circuit Status AVP indicating "inactive". On reception of this message, the ingress PE MUST generate AIS on the related ATM VPCs or VCCs.

Note: Since the Circuit Status AVP cannot distinguish between forward and backward failures, this procedure will result in AIS messages in both directions, even if the failure affected only one direction. An alternative procedure is the following. When an ingress PE detects a

failure, it sends an SLI message to its peer AND it generates an RDI message on the affected AC in the reverse direction. When the egress PE receives RDI from the far end of the affected segment, it will not send an SLI message to the ingress PE and ignore the RDI signal.

In case of transparent mapping, where the PE does not know which VCCs and/or VPCs are active, the near-end PE MUST send a PW-STATUS message to its peer. How the peer propagates that message on its AC is beyond the scope of this document.

7. SONET Encapsulation (CEP)

[CEP] discusses how Loss of Connectivity and other SONET/SDH protocol failures on the PW are translated to alarms on the ACs and vice versa. In essence, all fault management procedures are handled entirely in the emulated protocol. There is no need for an interaction between PW fault management and SONET layer fault management.

8. TDM Encapsulation

<TBD>

9. Ethernet Encapsulation

At this point in time, Ethernet OAM is not defined. Therefore, the procedures for mapping PW failures to Ethernet OAM messages and vice versa are currently rudimentary.

When an ingress PE detects that an Ethernet AC is down (because the related ethernet physical interface is down), it SHOULD send a PW Status message indicating both "AC Receive Fault" and "AC Transmit Fault".

If an egress PE determines that all ACs on a specific ethernet physical interface are affected (either because of ingress AC failures or because of PW failures), it MAY propagate these alarms by bringing the entire physical interface down.

10. Security Considerations

The mapping messages described in this document do not change the security functions inherent in the actual messages.

11. Acknowledgments

Hari Rakotoranto, Eric Rosen, Mark Townsley, Michel Khouderchah, Bertrand Duvivier, Vanson Lim and Chris Metz Cisco Systems

12. References

[BFD] Katz, D., Ward, D., "Bidirectional Forwarding Detection", Internet Draft <draft-katz-ward-bfd-01.txt>, August 2003

[CEP] Malis, A., et.al., "SONET/SDH Circuit Emulation over Packet (CEP)", Internet Draft <draft-ietf-pwe3-sonet-03.txt>, October 2003

[CONGESTION] Rosen, E., Bryant, S., Davie, B., "PWE3 Congestion Control Framework", Internet Draft <draft-rosen-pwe3-congestion-00.txt", October 2003

[CONTROL] Martini, L., Rosen, E., Smith, T., "Pseudowire Setup and Maintenance using LDP", Internet Draft <draft-ietf-pwe3-control-protocol-05.txt>, December 2003

[ICMP] Postel, J. "Internet Control Message Protocol" RFC 792

[ITU-T] Recommendation I.610 "B-ISDN operation and maintenance principles and functions", February 1999

[ITU-T] Recommendation I.620 "Frame relay operation and maintenance principles and functions", October 1996

[ITU-T] Recommendation Q.933 " ISDN Digital Subscriber Signalling System No. 1 (DSS1) û Signalling specifications for frame mode switched and permanent virtual connection control and status monitoring" February 2003

[ITU-T] Recommendation Y.1710 "Requirements for Operation & Maintenance functionality for MPLS networks", November 2002

[L2TPv3] Lau, J., et.al. " Layer Two Tunneling Protocol (Version 3", Internet Draft <draft-ietf-l2tpext-l2tp-base-11.txt>, October 2003

[LSPPING] Kompella, K., Pan, P., Sheth, N., Cooper, D., Swallow, G., Wadhwa, S., Bonica, R., " Detecting MPLS Data Plane Failures", Internet Draft < draft-ietf-mpls-lsp-ping-04.txt>, October 2003

[OAM REQ] T. Nadeau et.al., "OAM Requirements for MPLS Networks", Internet Draft <draft-ietf-mpls-oam-requirements-02>, June 2003

[PWEARCH] Bryant, S., Pate, P., "PWE3 Architecture", Internet Draft, < draft-ietf-pwe3-arch-06.txt>, October 2003

[PWEATM] Martini, L., et al., "Encapsulation Methods for Transport of ATM Cells/Frame Over IP and MPLS Networks", Internet Draft <draft-ietf-pwe3-atm-encap-03.txt>, October 2003

[PWREQ] Xiao, X., McPherson, D., Pate, P., "Requirements for Pseudo Wire Emulation Edge to-Edge (PWE3)", < draft-ietf-pwe3-requirements-08.txt>, December 2003

[RSVP-TE] Awduche, D., et.al. " RSVP-TE: Extensions to RSVP for LSP Tunnels", RFC 3209

[VCCV] Nadeau, T., et al."Pseudo Wire Virtual Circuit Connection Verification (VCCV)", Internet Draft <draft-ietf-pwe3-vccv-01.txt>, October 2003.

13. Intellectual Property Rights Notices

14. Full Copyright Statement

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.

Author's Addresses

Authors' Addresses

Thomas D. Nadeau
Cisco Systems, Inc.
250 Apollo Drive
Chelmsford, MA 01824
Email: tnadeau@cisco.com

Monique Morrow
Cisco Systems, Inc.
Glatt-com
CH-8301 Glattzentrum
Switzerland
Email: mmorrow@cisco.com

Yuichi Ikejiri
NTT Communications Corporation
1-1-6, Uchisaiwai-cho, Chiyoda-ku
Tokyo 100-8019
JAPAN
Email: y.ikejiri@ntt.com

Kenji Kumaki
KDDI Corporation
KDDI Bldg. 2-3-2
Nishishinjuku, Shinjuku-ku
Tokyo 163-8003
JAPAN
E-mail : ke-kumaki@kddi.com

Satoru Matsushima
Japan Telecom
JAPAN
Email: satoru@ft.solteria.net

Peter B. Busschbach
Lucent Technologies
67 Whippany Road
Whippany, NJ, 07981
Email: busschbach@lucent.com