Pseudo-Wire Edge-to-Edge(PWE3)
Internet Draft
Expiration Date: August 2005

Thomas D. Nadeau
Monique Morrow
Cisco Systems

Peter Busschbach
Lucent Technologies

Dave Allan
Nortel Networks

Mustapha Aissaoui
Alcatel

Editors

February 2005

Pseudo Wire (PW) OAM Message Mapping
draft-ietf-pwe3-oam-msg-map-02.txt

Status of this Memo

Abstract

This document specifies the mapping of defect states between a Pseudo Wire and the Attachment Circuits (AC) of the end-to-end emulated service. This document covers the case whereby the ACs and the PWs are of the same type in accordance to the PWE3 architecture [PWEARCH] such that a homogenous PW service can be constructed.

Table of Contents

1 Conventions used in this document

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

2 Contributors

Thomas D. Nadeau, tnadeau@cisco.com

Monique Morrow, mmorrow@cisco.com

Peter B. Busschbach, busschbach@lucent.com

Mustapha Aissaoui, mustapha.aissaoui@alcatel.com

Matthew Bocci, matthew.bocci@alcatel.co.uk

David Watkinson, david.watkinson@alcatel.com

Yuichi Ikejiri, y.ikejiri@ntt.com

Kenji Kumaki, kekumaki@kddi.com

Satoru Matsushima, satoru@ft.solteria.net

David Allan, dallan@nortelnetworks.com

Himanshu Shah, hshah@ciena.com

Simon Delord, simon.delord@francetelecom.com

3 Scope

This document specifies the mapping of defect states between a Pseudo Wire and the Attachment Circuits (AC) of the end-to-end emulated service. This document covers the case whereby the ACs and the PWs are of the same type in accordance to the PWE3 architecture [PWEARCH] such that a homogenous PW service can be constructed.

Ideally only PW and AC defects need be propagated into the Native Service (NS), and NS OAM mechanisms are transported transparently over the PW. Some homogenous scenarios use PW specific OAM mechanisms to synchronize defect state between PEs due to

discontinuities in native service OAM between the AC and the PW (e.g. FR LMI), or lack of native service OAM (e.g. Ethernet).

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.

This document covers PWE over MPLS PSN, PWE over IP PSN and PWE over L2TP PSN.

4 Terminology

AIS	Alarm Indication Signal
AC	Attachment circuit
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
NS Native Service
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:

The PW can ride over three types of Packet Switched Network (PSN). A PSN which makes use of LSPs as the tunneling technology to forward the PW packets will be referred to as an MPLS PSN. A PSN which makes use of MPLS-in-IP tunneling [MPLS-in-IP], with a MPLS shim header used as PW demultiplexer, will be referred to as an MPLS-IP PSN. A PSN, which makes use of L2TPv3 [L2TPv3] as the tunneling technology, will be referred to as L2TP-IP PSN.

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.

In the context of this document a PE forwards packets between an AC and a PW. The other PE that terminates the PW is the ææpeerÆÆ PE and the attachment circuit associated with the far end PW termination is the ææremote ACÆÆ.

Defects are discussed in the context of defect states, and the criteria to enter and exit the defect state.

The direction of defects is discussed from the perspective of the observing PE and what the PE may explicitly know about information transfer capabilities of the PW service.

A forward defect is one that impacts information transfer to the observing PE. It impacts the observing PEÆs ability to receive information. A forward defect MAY also imply impact on information

sent or relayed by the observer (and as it cannot receive is therefore unknowable) and so the forward defect state is considered to be a superset of the two defect states.

A reverse defect is one that uniquely impacts information sent or relayed by observer.

At the present time code points for forward defect and reverse defect have not been specified for BFD and LDP PW control. These are referred to as ææforward defectÆÆ and æereverse defectÆÆ indications as placeholders for code point assignment. However, a mapping to existing PW status code points [IANA] may be performed:

```
Forward defect - corresponds to the logical OR of
        Local Attachment Circuit ( ingress ) Receive Fault
                        AND
        Local PSN-facing PW ( egress ) Transmit Fault

Reverse defect - corresponds to the logical OR of
        Local Attachment Circuit ( egress ) Transmit Fault
                        AND
        Local PSN-facing PW ( egress ) Receive Fault
```

5 Reference Model and Defect Locations

Figure 1 illustrates the PWE3 network reference model with an indication of the possible defect locations. This model will be referenced in the remainder of this document for describing the OAM procedures.

```
             ACs             PSN tunnel               ACs
                    +----+                  +----+
+----+              | PE1|==================| PE2|          +----+
|    |---(a)---(b)..(c)......PW1..(d)..(c)..(f)---(e)---|    |
| CE1|   (N1)       |    |                  |    |   (N2)   |CE2 |
|    |----------|............PW2.............|----------|    |
+----+              |    |==================|    |          +----+
     ^              +----+                  +----+          ^
```



```
     |      Provider Edge 1         Provider Edge 2      |
     |                                                   |
     |<-------------- Emulated Service ---------------->|
  Customer                                            Customer
   Edge 1                                              Edge 2
```

Figure 1: PWE3 Network Defect Locations

In all interworking scenarios described in this document, it is assumed that at PE1 the AC and the PW are of the same type. The procedures described in this document exclusively apply to PE1. PE2 for a homogenous service implements the identical functionality (although it is not required to as long as the notifications across the PWs are consistent).

The following is a brief description of the defect locations:

(a) Defect in the first L2 network (N1). This covers any defect in the N1 which impacts all or a subset of ACs terminating in PE1. The defect is conveyed to PE1 and to the remote L2 network (N2) using the native service specific OAM defect indication.
(b) Defect on a PE1 AC interface.
(c) Defect on a PE PSN interface.
(d) Defect in the PSN network. This covers any defect in the PSN which impacts all or a subset of the PSN tunnels and PWs terminating in a PE. The defect is conveyed to the PE using a PSN and/or a PW specific OAM defect indication. Note that control plane, i.e., signaling and routing, messages do not necessarily follow the path of the user plane messages. Defect in the control plane are detected and conveyed separately through control plane mechanisms. However, in some cases, they have an impact on the status of the PW as explained in the next section.
(e) Defect in the second L2 network (N2). This covers any defect in N2 which impacts all or a subset of ACs terminating in PE2 (which is considered a ææremote AC defectÆÆ in the context of procedures outlined in this draft). The defect is conveyed to PE2 and to the remote L2 network (N1) using the native service OAM defect indication.
(f) Defect on a PE2 AC interface (which is also considered a ææremote AC defectÆÆ in the context of this draft).

6 Abstract Defect States

PE1 is obliged to track four abstract defect states that reflect the observed state of both directions of the PW service on both the AC and the PW sides. Faults may impact only one or both directions of the PW.

The observed state is a combination of faults directly detected by PE1, or faults it has been made aware of via notifications.

```
                              +------+
         ----AC forward---->|      |-----PW reverse---->
   CE1                        | PE1 |                        PE2/CE2
         <---AC reverse-----|      |<----PW forward-----
                              +------+
```

(arrows indicate direction of traffic)

Figure 2: Forward and Reverse Defect States

PE1 will directly detect or be notified of AC forward and PW forward defects as they occur upstream of PE1 and impact traffic being sent to PE1. PE1 will only be notified of AC reverse and PW reverse defects as they universally will be detected by other devices and only impact traffic that has already been relayed by PE1.

The procedures outlined in this document define the entry and exit criteria for each of the four states with respect to the set of potential ACs and PWs within the document scope and the consequent actions that PE1 must perform to properly interwork those notifications. The abstract defect states used by PE1 are common to all potential interworking combinations of PWs and ACs.

When a PE has multiple sources of notifications from a peer (e.g. PSN and LDP control plane), it is obliged to track all sources, but with respect to consequent actions the forward state ALWAYS has precedence over the reverse state.

7 PW Status and Defects

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

7.1 PW Defects

Possible defects that impact PWs are the following.

. Physical layer defect in the PSN interface

. PSN tunnel failure which results in a loss of connectivity between ingress and egress PE.

. Control session failures between ingress and egress PE

In case of an MPLS PSN and an MPLS-IP PSN there are additional defects:

. PW labeling error, which is due to a defect in the ingress PE,or to an over-writing of the PW label value somewhere along the LSP

path.

. LSP tunnel Label swapping errors or LSP tunnel label merging errors in the MPLS network. This 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.

7.1.1 Packet Loss

Persistent congestion in the PSN or in a PE could impact the proper operation of the emulated service.

A PE can detect packet loss resulting from congestion through several methods. If a PE uses the sequence number field in the PWE3 Control Word for a specific Pseudo Wire [PWEARCH], it has the ability to detect packet loss. [CONGESTION] discusses other possible mechanisms to detect congestion between PWs.

Generally, there are congestion alarms which are raised in the node and to the management system when congestion occurs. The decision to declare the PW Down and to re-signal it through another path is usually at the discretion of the network operator.

7.2 Defect Detection and Notification

7.2.1 Defect Detection Tools

To detect the defects listed in 7.1, Service Providers have a variety of options available:

Physical Layer defect detection and notification mechanisms such as SONET/SDH LOS, LOF,and AIS/FERF.

PSN Defect Detection Mechanisms:

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

For PWE3 over an MPLS PSN and an MPLS-IP PSN, several tools can be

used.
. LSP-Ping and LSP-Traceroute([LSPPING]) for LSP tunnel connectivity verification.

. LSP-Ping with Bi-directional Forwarding Detection ([BFD]) for LSP tunnel continuity checking.

.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]), but only at the control plane.

PW specific defect detection mechanisms:

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

Furthermore, the detection of a fault could occur at different points in the network and there are several ways the observing PE determines a fault exists:

a. egress PE detection of failure (e.g. BFD)
b. ingress PE detection of failure (e.g. LSP-PING)
c. ingress PE notification of failure (e.g. RSVP Path-err)

7.2.2 Defect Detection Mechanism 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 defect detection, as opposed to diagnostic usage, they must be run periodically.

. Control protocol failure indications, e.g. detected through L2TP Keep-alive messages or the RSVP-TE Hello messages, can be used to detect many network failures. However, control protocol failures

do not necessarily coincide with data plane failures. Therefore, a defect detection mechanism in the data plane is required to protect against all potential data plane failures. Furthermore, fault diagnosis mechanisms for data plane failures are required to further analyze detected failures.

. For PWE3 over an MPLS PSN and an MPLS-IP PSN, it is effective to run a defect detection mechanism over a PSN Tunnel frequently and run one over every individual PW within that PSN Tunnel less frequently. 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 defect 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 and an MPLS-IP PSN, PEs could detect PSN label errors through the execution of LSP-Ping. However, use of VCCV is preferred as it is a more accurate detection tool for pseudowires.
Furthermore, it can be run using a BFD mode, i.e., VCCV-BFD, which allows it to be used as a light-weight detection mechanism for PWs. 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. LSP ping and/or LSP trace could then be used to diagnose the detected failure.

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.

7.3 Overview of fault notifications

For a MPLS PSN and a IP PSN using MPLS-in-IP (MPLS-IP PSN), PW status signaling messages are used as the default mechanism for AC and PW status and defect indication [PWE3-CONTROL].

For a IP PSN using L2TPv3, i.e., a L2TP-IP PSN, StopCCN and CDN messages are used for conveying defects in the PSN and PW

respectively, while the Set-Link-Info (SLI) messages are used to convey status and defects in the AC and local L2 network.

Optionally, PEs can negotiate the use of VCCV-BFD for both PW fault detection and AC/PW fault notifications as explained in [VCCV]. What BFD is used for is negotiated:

i. not used
ii. used for PW fault detection (which implies reverse notifications)
iii. used for PW fault detection and all PW/AC fault notifications

When BFD is to be used for all fault notifications, then BFD is the preferred mechanism of exchanging fault notifications.

PE1 will translate the PW defect states 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.

7.3.1 Use of Native Service notifications

In the context of this document, ATM and unstructured SONET/TDM PWs are the only examples of a PW that has native service notification capability. Frame relay does have the FR OAM specification [FRF.19], but this is not commonly deployed. All other PWs use PW specific notification mechanisms.

ATM PWs may optionally also use PW specific notification mechanisms.

In normal, i.e., defect-free, operation, all the types of ATM OAM cells described in Section 14.2 are either terminated at the PE, for OAM segments terminating in the AC endpoint, or transparently carried over the PSN tunnel [PWE3-ATM]. This is referred to as ææinband ATM OAM over PWÆÆ and is the default method.

An optional out-of band method based on relaying the ATM defect state over a PW specific defect indication mechanism is provided for PEÆs which cannot generate and/or transmit ATM OAM cells over the ATM PW. This is referred to as ææOut-of-band ATM OAM over PWÆÆ.

7.3.2 The Use of PW Status for MPLS and MPLS-IP PSNs

This document specifies the use of PW status signaling as the default mechanism for the purpose of conveying the status of a PW

and ACs between PEs.

For a MPLS PSN and a IP PSN using MPLS-in-IP (MPLS-IP PSN), PW status signaling messages are used as the default mechanism for AC and PW status and defect indication [PWE3-CONTROL].

PW status is used to convey the defect view of the PW local to the originating PE. This is the local PW state, and when the NS does not have native OAM capability or emulation of native capability is prohibitive, the AC state. This is in the form of a ææforward defectÆÆ or a æærevers defectÆÆ.

7.3.3 The Use of L2TP STOPCCN and CDN

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

. Failures detected through defect detection mechanisms in the L2TP-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.

7.3.4 The Use of BFD Diagnostic Codes

If the PEs have negotiated the use of VCCV-BFD for both PW fault detection and AC/PW fault notifications as explained in [VCCV]

then BFD is the preferred mechanism of exchanging fault notifications.

[BFD] defines a set of diagnostic codes that partially overlap with failures that can be communicated through PW Status messages or L2TP STOPCCN and CDN messages. 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:

Code	Message
0	No Diagnostic
1	Control Detection Time Expired
3	Neighbor Signaled Session Down
7	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 STOPCCN 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.

In conclusion, one would communicate PW defects through PW Status messages, or L2TP STOPCCN and CDN messages in all cases, except for a well-defined set of exceptions where BFD is used. How PW defects that can be detected through the use of BFD or through other means, are mapped to defect indications on the ACs is described in section Error! Reference source not found. and in subsequent sections.

8 PW Defect State Entry/Exit

8.1 PW Forward Defect Entry/Exit

A PE will enter the PW forward defect state if one of the following occurs

. It detects loss of connectivity on the PSN tunnel over which the PW is riding. This includes label swapping errors and label merging errors.

. It receives a message from PE2 indicating PW ææforward defectÆÆ or ææPW not forwardingÆÆ, which indicates PE2 detected or was notified of a PW fault downstream of it or that there was a remote AC fault.

In the case of an L2TP-IP, this is a L2TP StopCCN or CDN message. A StopCCN message indicates that the control connection has been shut down by the remote PE [L2TPv3]. This is typically used for defects in the PSN which impact both the co
ntrol connection and the individual data plane sessions. On reception of this message, a PE closes the control connection and will clear all the sessions managed by this control connection. Since each session carries a single PW, the state of the corresponding PWs is changed to DOWN. A CDN message indicates that the remote peer requests the disconnection of a specific session [L2TPv3]. In this case only the state of the corresponding PW is changed to DOWN. This is typically used for local defects in a PE which impact only a specific session and the corresponding PW.

. It detects a loss of PW connectivity, including label errors, through VCCV-BFD or VCCV-PING in no reply mode.

Note that if the PW control session between the PEs fails, the PW is torn down and needs to be re-established. However, the consequent actions towards the ACs are the same as if the PW entered the forward defect state. Precise details of AC defect state entry and exit criteria are specified elsewhere (e.g. I.610) and such references will supersede the descriptions herein.

PE1 will exit the forward defect state if the notified PW status from the PE2 has the ææforward defectÆÆ indication clear, and it has established that PW/PSN connectivity is working in the forward direction. Note that this may result in a transition to the PW working or PW reverse defect states.
For a PWE3 over a L2TP-IP PSN, a PE will exit the PW forward defect state when the following conditions are true:

. All defects it had previously detected have disappeared, and

. A L2TPv3 session is successfully established to carry the PW packets.

8.2 PW reverse defect state entry/exit

A PE will enter the PW reverse defect state if one of the following occurs

. It receives a message from PE2 indicating PW ææreverse defectÆÆ which indicates PE2 detected or was notified of a PW/PSN fault upstream of it or that there was a remote AC fault and it is not already in the PW forward defect state.

PE1 will exit the reverse defect state if the notified PW status from the PE2 has the ææreverse defectÆÆ indication clear, or it has entered the PW forward defect state.

For a PWE3 over a L2TP-IP PSN, a PE will exit the PW reverse defect state when the following conditions are true:

. All defects it had previously detected have disappeared, and

. A L2TPv3 session is successfully established to carry the PW packets.

8.2.1 PW reverse defects that are treated as AC Forward Defects

Some PW mechanisms will result in PW defects being detected by or notified to PE1 when PE1 is upstream of the fault but the notification did not originate with PE2. The resultant actions are identical to that of entering the AC forward defect state as PE1 needs to synchronize state with PE2 and the PW state communicated from PE1 to PE2 needs to indicate state accordingly.

When the PSN uses RSVP-TE or proactively uses LSP-PING as a PW fault detection mechanism, PE1 must consider entry to the AC forward defect state to be the logical or of the AC entry criteria outlined for each AC type in the subsequent sections, and that of the known PW state in the direction of PE2 downstream of PE1 (indicated via RSVP patherr or LSP-PINGs).

The exit criteria being when the logical AND of the RSVP fault state, LSP-PING fault state and the actual AC forward defect exit

criteria has been met, indicating no forward defects.

9 AC Defect States

9.1 FR ACs

PE1 enters the AC Forward Defect state if any of the following conditions are met:

(i) A PVC is not æddeletedÆ from the Frame Relay network and the Frame Relay network explicitly indicates in a full status report (and optionally by the asynchronous status message) that this Frame Relay PVC is æinactiveÆ. In this case, this status maps across the PE to the corresponding PW only.

(ii) The LIV indicates that the link from the PE to the Frame Relay network is down. In this case, the link down indication maps across the PE to all corresponding PWs.

(iii) A physical layer alarm is detected on the FR interface. In this case, this status maps across the PE to all corresponding PWs.

A PE exits the AC Forward Defect state when all defects it had previously detected have disappeared.

The AC reverse defect state is not valid for FR ACs.

9.2 ATM ACs

9.2.1 AC Forward Defect State Entry/Exit

PE1 enters the AC forward defect state if any of the following conditions are met:

(i) It detects or is notified of a physical layer fault on the ATM interface and/or it terminates an F4 AIS flow or has loss of F4 CC for a VP carrying VCCÆs.

(ii) It terminates an F4/F5 AIS OAM flow indicating that the ATM VP/VC is down in the adjacent L2 ATM network (e.g., N1 for PE1). This is applicable to the case of the ææout-of-band ATM OAM over PWÆÆ method only.

(iii) It detects loss of connectivity on the NS ATM VPC/VCC while terminating ATM continuity checking (ATM CC) with the local ATM network and CE.

A PE exits the AC Forward Defect state when all defects it had

previously detected have disappeared. The exact conditions under which a PE exits the AIS state, or declares that connectivity is restored via ATM CC are defined in I.610 [I.610].

9.2.2 AC Reverse Defect State Entry/Exit

A PE enters the AC reverse defect state if any of the following conditions are met:

(i) It terminates an F4/F5 RDI OAM flow indicating that the ATM VP/VC AC is down in the adjacent L2 ATM network (e.g., N1 for PE1). This is applicable to the case of out-of-band ATM OAM over PW only.

A PE exits the AC Reverse Defect state if the AC state transitions to working or to the AC forward defect state. The criteria for exiting the RDI state are described in I.610.

9.3 Ethernet AC State

PE1 enters the forward defect state if any of the following conditions are met:

(i) A physical layer alarm is detected on the Ethernet interface.

A PE exits the Ethernet AC forward defect state when all defects it had previously detected have disappeared.

10 PW Forward Defect Entry/Exit procedures

10.1 PW Forward Defect Entry Procedures

10.1.1 FR AC procedures

These procedures are applicable only if the transition from the working state to the PW Forward defect state. A transition from PW reverse defect state to the forward defect state does not require any additional notification procedures to the FR AC as it has already been told the peer is down.

(i) PE1 MUST generate a full status report with the Active bit = 0 (and optionally in the asynchronous status message), as per Q.933 annex A, into N1 for the corresponding FR ACs.

10.1.2 Ethernet AC Procedures

No procedures are currently defined.

10.1.3 ATM AC procedures

On entry to the PW Forward Defect State

(i) PE1 MUST commence F5 AIS insertion into the corresponding AC.

(ii) PE1 MUST terminate any F5 CC generation on the corresponding AC.

10.1.4 Additional procedures for a FR PW, an ATM PW in the ææout-of-band ATM OAM over PW methodÆÆ, and an Ethernet PW

If the PW failure was explicitly detected by PE1, it MUST assume PE2 has no knowledge of the defect and MUST notify PE2 in the form of a reverse defect notification:

For PW over MPLS PSN or MPLS-IP PSN

(i) A PW Status message indicating a ææreverse defectÆÆ, or

(ii) A VCCV-BFD diagnostic code if the optional use of VCCV-BFD notification has been negotiated

For PW over L2TP-IP PSN

(i) An L2TP Set-Link Info (LSI) message with a Circuit Status AVP indicating "active" Or,

(ii) A VCCV-BFD diagnostic code if the optional use of VCCV-BFD notification has been negotiated

Otherwise the entry to the defect state was the result of a notification from PE2 (indicating that PE2 already had knowledge of the fault) or loss of the control adjacency (similarly visible to PE2).

10.2 PW Forward Defect Exit Procedures

10.2.1 FR AC procedures

On transition from the PW forward defect state to the reverse defect state PE1 takes no action w.r.t. the AC.

On exit from the PW Forward defect state

(i) PE1 MUST generate a full status report with the Active bit = 1 (and optionally in the asynchronous status message), as per Q.933 annex A, into N1 for the corresponding FR ACs.

10.2.2 Ethernet AC Procedures

No procedures are currently defined

10.2.3 ATM AC procedures

On exit from the PW Forward Defect State

(i) PE1 MUST cease F5 AIS insertion into the corresponding AC.

(ii) PE1 MUST resume any F5 CC generation on the corresponding AC.

10.2.4 Additional procedures for a FR PW, an ATM PW in the ææout-of-band ATM OAM over PWÆÆ method, and an Ethernet PW

If the PW failure was explicitly detected by PE1, it MUST notify PE2 in the form of clearing the reverse defect notification:

For PW over MPLS PSN or MPLS-IP PSN

(i) A PW Status message with the ææreverse defectÆÆ indication clear, and the remaining indicators showing either working or a transition to the ææforward defectÆÆ state. Or,

(ii) A VCCV-BFD diagnostic code with the same attribute as (i) if the optional use of VCCV-BFD notification has been negotiated

For PW over L2TP-IP PSN

(i) An L2TP Set-Link Info (LSI) message with a Circuit Status AVP indicating "active" Or,

(ii) A VCCV-BFD diagnostic code with the same attributes as (i) if the optional use of VCCV-BFD notification has been negotiated

10.3 PW Reverse Defect Entry Procedures

10.3.1 FR AC procedures

On transition from the PW forward defect state to the reverse defect state PE1 takes no action w.r.t. the AC.

On entry to the PW reverse defect state

(i) PE1 MUST generate a full status report with the Active bit = 0 (and optionally in the asynchronous status message), as per Q.933 annex A, into N1 for the corresponding FR ACs.

10.3.2 Ethernet AC Procedures

No procedures are currently defined

10.3.3 ATM AC procedures

On entry to the PW Reverse Defect State

(i) PE1 MUST commence F5 RDI insertion into the corresponding AC. This applies to the case of an ATM PW in the ææout-of-band ATM OAM over PWÆÆ method only.

10.4 PW Reverse Defect Exit Procedures

10.4.1 FR AC procedures

On transition from the PW reverse defect state to the PW forward defect state PE1 takes no action with respect to the AC.

On exit from the PW Reverse defect state

(i) PE1 MUST generate a full status report with the Active bit = 1 (and optionally in the asynchronous status message), as per Q.933 annex A, into N1 for the corresponding FR ACs.

10.4.2 Ethernet AC Procedures

No procedures are currently defined

10.4.3 ATM AC procedures

On exit from the PW Reverse Defect State

(i) PE1 MUST cease F5 RDI insertion into the corresponding AC. This applies to the case of an ATM PW in the ææout-of-band ATM OAM over PWÆÆ method only.

10.5 Procedures in FR Port Mode

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 of the Q.933 link and PVC management protocol

at the Frame Relay devices at one or both sites of the emulated interface.

10.6 Procedures in ATM Port Mode

In case of transparent cell transport, i.e., "port mode", where the PE does not keep track of the status of individual ATM VPCs or VCCs, a PE cannot relay PW defect state over these VCCs and VPCs. If ATM CC is run on the VCCs and VPCs end-to-end (CE1 to CE2), or on a segment originating and terminating in the ATM network and

spanning the PSN network, it will timeout and cause the CE or ATM switch to enter the ATM AIS state.

11 AC Defect Entry/Exit Procedures

11.1 AC Forward defect entry:

On entry to the forward defect state, PE1 may need to perform procedures on both the PW and the AC.

11.1.1 Procedures for a FR PW, an ATM PW in the ææout-of-band ATM OAM over PWÆÆ method, or an Ethernet PW

On entry to the AC forward defect state, PE1 notifies PE2 of a forward defect:

For PW over MPLS PSN or MPLS-IP PSN

(i) A PW Status message indicating ææforward defectÆÆ, or

(ii) A VCCV-BFD diagnostic code of ææforward defectÆÆ if the optional use of VCCV-BFD notification has been negotiated.

For PW over L2TP-IP PSN

(i) An L2TP Set-Link Info (LSI) message with a Circuit Status AVP indicating "inactive", or

(ii) A VCCV-BFD diagnostic code of ææforward defectÆÆ if the optional use of VCCV-BFD notification has been negotiated.

11.1.2 Procedures for a ATM PW in the ææinband ATM OAM over PWÆÆ method

On entry to the AC forward defect state, PE1 MUST:

a. Commence insertion of ATM AIS cells into the corresponding PW.

b. If PE1 is originating F4 or F5 I.610 CC cells, PE1 will suspend CC generation for the duration of the defect state.

11.1.3 Additional procedures for ATM ACs

On entry to the AC forward defect state PE1 will commence RDI insertion into the AC as per I.610. This procedure is applicable to the ææout-of-band ATM OAM over PWÆÆ method only.

11.2 AC Reverse defect entry

11.2.1 Procedures for a FR PW, an ATM PW in the ææout-of-band ATM OAM over PWÆÆ method, or an Ethernet PW

On entry to the AC reverse defect state, PE1 notifies PE2 of a reverse defect:

For PW over MPLS PSN or MPLS-IP PSN
(iii) A PW Status message indicating ææreverse defectÆÆ,or
(iv) A VCCV-BFD diagnostic code of ææreverse defectÆÆ if the optional use of VCCV-BFD notification has been negotiated.

For PW over L2TP-IP PSN
(iii) An L2TP Set-Link Info (LSI) message with a Circuit Status AVP indicating "inactive", or
(iv) A VCCV-BFD diagnostic code of ææreverse defectÆÆ if the optional use of VCCV-BFD notification has been negotiated.

11.2.2 Procedures for a ATM PW in the ææinband ATM OAM over PWÆÆ method

There are no procedures in this case as the AC reverse defect state is not valid for PE1 operating in this method.

11.3 AC Forward Defect Exit

11.3.1 Procedures for a FR PW, an ATM PW in the ææout-of-band ATM OAM over PWÆÆ method, or an Ethernet PW

On exit from the AC forward defect state PE1 notifies PE2 that the forward defect state has cleared (note that this may be a direct state transition to either the working state or the reverse defect state):

For PW over MPLS PSN or MPLS-IP PSN
(i) A PW Status message with forward defect clear and the remaining indicators showing either working or reverse defect state, or
(ii) A VCCV-BFD diagnostic code with the same attributes as (i) if the optional use of VCCV-BFD notification has been negotiated.

For PW over L2TP-IP PSN
(i) An L2TP Set-Link Info (LSI) message with a Circuit Status AVP indicating "active", or
(ii) A VCCV-BFD diagnostic code with the same attributes as (i) if the optional use of VCCV-BFD notification has been negotiated.

11.3.2 Procedures for a ATM PW in the ææinband ATM OAM over PWÆÆ method

On exit from the AC forward defect state, PE1 MUST:

(i) Cease insertion of ATM AIS cells into the corresponding PW.

(ii) If PE1 is originating F4 or F5 I.610 CC cells, PE1 will resume CC generation for the duration of the defect state.

11.3.3 Additional procedures for ATM ACs

On exit from the AC forward defect state PE1 will cease RDI insertion into the AC as per I.610. This procedure is applicable to the ææout-of-band ATM OAM over PWÆÆ method only.

11.4 AC Reverse Defect Exit

11.4.1 Procedures for a FR PW, an ATM PW in the ææout-of-band ATM OAM over PWÆÆ method, or an Ethernet PW

On exit from the AC reverse defect state, PE1 notifies PE2 that the reverse defect state has cleared (note that this may be a direct state transition to either the working state or the forward defect state):

For PW over MPLS PSN or MPLS-IP PSN

(i) A PW Status message with the ææreverse defectÆÆ indicator cleared and the remaining indicators showing either working or a transition to the ææforward defectÆÆ state, or

(ii) A VCCV-BFD diagnostic code with the same information as (i) if the optional use of VCCV-BFD notification has been negotiated.

For PW over L2TP-IP PSN

(i) An L2TP Set-Link Info (LSI) message with a Circuit Status AVP indicating "active", or

(ii) A VCCV-BFD diagnostic code with the same information as (i) if the optional use of VCCV-BFD notification has been negotiated.

11.4.2 Procedures for a ATM PW in the ææinband ATM OAM over PWÆÆ method

There are no procedures in this case as the AC reverse defect state is not valid for PE1 operating in this method.

12 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 defect management procedures are handled entirely in the emulated protocol. There is no need for an

interaction between PW defect management and SONET layer defect management.

13 TDM Encapsulation

From an OAM perspective, the PSN carrying a TDM PW provides the same function as that of SONET/SDH or ATM network carrying the same low-rate TDM stream. Hence the interworking of defect OAM is similar.

For structure-agnostic TDM PWs, the TDM stream is to be carried transparently across the PSN, and this requires TDM OAM indications to be transparently transferred along with the TDM data. For structure-aware TDM PWs the TDM structure alignment is terminated at ingress to the PSN and regenerated at egress, and hence OAM indications may need to be signaled by special means. In both cases generation of the appropriate emulated OAM indication may be required when the PSN is at fault.

Since TDM is a real-time signal, defect indications and performance measurements may be classified into two classes, urgent and deferrable. Urgent messages are those whose contents may not be significantly delayed with respect to the TDM data that they potentially impact, while deferrable messages may arrive at the far end delayed with respect to simultaneously generated TDM data. For example, a forward indication signifying that the TDM data is invalid (e.g. TDM loss of signal, or MPLS loss of packets) is only of use when received before the TDM data is to be played out towards the far end TDM system. It is hence classified as an urgent message, and we can not delegate its signaling to a separate maintenance or management flow. On the other hand, the forward loss of multiframe synchronization, and most reverse indications do not need to be acted upon before a particular TDM frame is played out.

From the above discussion it is evident that the complete solution to OAM for TDM PWs needs to have at least two, and perhaps three components. The required functionality is transparent transfer of native TDM OAM and urgent transfer of indications (by flags) along with the impacted packets. Optionally there may be mapping between TDM and PSN OAM flows.

TDM AIS generated in the TDM network due to a fault in that network is generally carried unaltered, although the TDM

encapsulations allow for its suppression for bandwidth conservation purposes. Similarly, when the TDM loss of signal is detected at the PE, it will generally emulate TDM AIS.

SAToP and the two structure-aware TDM encapsulations have converged on a common set of defect indication flags in the PW control word. When the PE detects or is informed of lack of validity of the TDM signal, it raises the local ("L") defect flag, uniquely identifying the defect as originating in the TDM network. The remote PE must ensure that TDM AIS is delivered to the remote TDM network. When the defect lies in the MPLS network, the remote

PE fails to receive packets. The remote PE generates TDM AIS towards its TDM network, and in addition raises the remote defect ("R") flag in its PSN-bound packets, uniquely identifying the defect as originating in the PSN. Finally, defects in the remote TDM network that cause RDI generation in that network, may optionally be indicated by proper setting of the field of valid packets in the opposite direction.

14 Appendix A: Native Service Management

14.1 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. TO achieve commonality with other technologies, æædownÆÆ is represented as a forward defect.

Frame relay connection management is not implemented over the PW using either of the techniques native to FR, therefore PW mechanisms are used to synchronize the view each PE has of the remote NS/AC. A PE will treat a remote NS/AC failure in the same way it would treat a PW or PSN failure, that is using AC facing FR connection management to notify the CE that FR is æædownÆÆ.

14.2 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 to instrument virtual paths and virtual channels respectively with regard to their performance and availability. 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.

14.3 Ethernet Management

At this point in time, inband Ethernet OAM standards are being specified in the International Telecommunications Union -
-
Telecommunications (ITU-T) and the Institute of Electrical and Electronics Engineers (IEEE). However, it will take some time

before they are widely deployed. Therefore, this document specifies only the procedures for mapping a defect due to a Ethernet physical layer fault. Defects on a remote Ethernet AC or defects in a PW cannot be mapped back to the local Ethernet network.

15 Security Considerations

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

16 Acknowledgments

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

17 References

[BFD] Katz, D., Ward, D., "Bidirectional Forwarding Detection", Internet Draft <draft-katz-ward-bfd-02.txt>, May 2004

[CEP] Malis, A., et.al., "SONET/SDH Circuit Emulation over Packet (CEP)", Internet Draft <draft-ietf-pwe3-sonet-09.txt>, August 2004

[CONGESTION] Rosen, E., Bryant, S., Davie, B., "PWE3 Congestion Control Framework", Internet Draft <draft-rosen-pwe3-congestion-02.txt", September 2004

[CONTROL] Martini, L., Rosen, E., Smith, T., "Pseudowire Setup and Maintenance using LDP", Internet Draft <draft-ietf-pwe3-control-protocol-14.txt>, December 2004

[FRF.19] Frame Relay Forum, ææFrame Relay Operations, Administration, and Maintenance Implementation AgreementÆÆ, March 2001.

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

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

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

[ITU-T Q.933] Recommendation Q.933 " ISDN Digital Subscriber Signalling System No. 1 (DSS1) 0 Signalling specifications for frame mode switched and permanent virtual connection control and status monitoring" February 2003

[L2TPv3] Lau, J., et.al. " Layer Two Tunneling Protocol (Version 3", Internet Draft <draft-ietf-l2tpext-l2tp-base-15.txt>, December 2004

[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-07.txt>, October 2004

[MPLS-in-IP] Worster. T., et al., ææEncapsulating MPLS in IP or

Generic Routing Encapsulation (GRE)ÆÆ, draft-ietf-mpls-in-ip-or-gre-08.txt, June 2004.

[OAM REQ] T. Nadeau et.al., "OAM Requirements for MPLS Networks", Internet Draft <draft-ietf-mpls-oam-requirements-05>, December 2004

[PWEARCH] Bryant, S., Pate, P., "PWE3 Architecture", Internet Draft, < draft-ietf-pwe3-arch-07.txt>, March 2004

[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-07.txt>, Ocotber 2004

[PWREQ] Xiao, X., McPherson, D., Pate, P., "Requirements for Pseudo Wire Emulation Edge to-Edge (PWE3)", RFC 3916, September 2004

[RSVP-TE] Awduche, D., et.al. " RSVP-TE: Extensions to RSVP for LSP Tunnels", RFC 3209, December 2001

[VCCV] Nadeau, T., et al."Pseudo Wire Virtual Circuit Connection Verification (VCCV)", Internet Draft <draft-ietf-pwe3-vccv-04.txt>, February 2005.

18 Intellectual Property Disclaimer

to practice this standard. Please address the information to the IETF Executive Director.

19 Full Copyright Statement

20 Authors' Addresses

Thomas D. Nadeau
Cisco Systems, Inc.
300 Beaverbrook Drive
Boxborough, MA 01824
Phone: +1-978-936-1470
Email: tnadeau@cisco.com

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

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

Mustapha Aissaoui
Alcatel
600 March Rd
Kanata, ON, Canada. K2K 2E6
Email: mustapha.aissaoui@alcatel.com

Matthew Bocci
Alcatel
Voyager Place, Shoppenhangers Rd
Maidenhead, Berks, UK SL6 2PJ
Email: matthew.bocci@alcatel.co.uk

David Watkinson
Alcatel
600 March Rd
Kanata, ON, Canada. K2K 2E6
Email: david.watkinson@alcatel.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 : kekumaki@kddi.com

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

David Allan
Nortel Networks
3500 Carling Ave.,
Ottawa, Ontario, CANADA
Email: dallan@nortelnetworks.com

Simon Delord
France Telecom
2 av, Pierre Marzin
22300 LANNION, France
E-mail: simon.delord@francetelecom.com