PWE3 Working Group
Internet Draft
Expiration Date: December 2002

Andrew G. Malis
Ken Hsu
Vivace Networks, Inc.

David Zelig
Corrigent Systems, LTD.

Jeremy Brayley
Steve Vogelsang
John Shirron
Laurel Networks, Inc.

Jim Boyle
Protocol Driven Networks, Inc.

Luca Martini
Craig White
Level 3 Communications, LLC.

Ron Cohen
Lycium Networks

Prayson Pate
Overture Networks, Inc.

Tom Johnson
Marlene Drost
Ed Hallman
Litchfield Communications, Inc.

June 2002

SONET/SDH Circuit Emulation over Packet (CEP)
draft-malis-pwe3-sonet-03.txt

Status of this Memo

Abstract

Generic requirements and framework for Pseudo Wire Emulation Edge-to-Edge (PWE3) have been described in [PWE3-REQ] and [PWE3-FW]. This draft provides encapsulation formats and semantics for connecting SONET/SDH edge networks through a packet network using IP or MPLS. This basic application of SONET/SDH interworking will allow service providers to take advantage of new technologies in the core in order to provide traditional SONET/SDH services.

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 Introduction

This document describes a protocol that performs SONET Emulation over a variety of Packet-Switched Networks (PSNs) as part of the PWE3 Working Group. The document assumes that the reader is familiar with the PWE3 terminology and concepts described in PWE3 requirements and framework documents [PWE3-REQ] and [PWE3-FW] as well as the PWE3 Protocol Layering Model [PWE3-LAYERS]. The protocol is titled "Circuit Emulation over Packet" (CEP).

The transmission system for circuit-oriented TDM signals is the

Synchronous Optical Network [SONET], [GR253] / Synchronous Digital Hierarchy (SDH) [G707]. To support TDM traffic (which includes voice, data, and private leased line services) PSNs must emulate the circuit characteristics of SONET/SDH payloads. An RTP Header [RFC1889] and a CEP Control Word are used to encapsulate the SONET/SDH TDM signals for transmission over an arbitrary PSN.

This document also describes an optional extension to CEP called Dynamic Bandwidth Allocation (DBA). This is a method for dynamically reducing the bandwidth utilized by emulated SONET/SDH circuits in the packet network. This bandwidth reduction is accomplished by not sending the SONET/SDH payload through the packet network under certain conditions such as AIS-P or STS SPE Unequipped.

In addition, this document describes a technique for RTP header compression/suppression based on [ROHC-LLA].

This document is based on a previous document describing a method for encapsulating SONET signals for carriage over MPLS networks [CEM].

This document is closely related to and references [MARTINI-TRANS], which describes the control protocol methods used to signal the usage of CEP, [MARTINI-ENCAP] which describes a related method of encapsulating Layer 2 frames over MPLS and which shares the same signaling, and [CEM-MIB] which describes a MIB for controlling and observing CEM services.

This document is complimentary to [CESoPSN] and [CEP-VT] which describe methods for transporting sub-STS-1 rate circuits in native format or VT mapped respectively.

3 Applicability Statement

SONET/SDH Circuit Emulation over Packet (CEP) is an encapsulation layer intended for emulating SONET/SDH circuits over a Packet Switched Network.

This protocol provides a method for emulating the key elements of traditional SONET/SDH SPE services across a packet-switched network. Both large fixed-facility network operators and smaller network operators using ad hoc facilities may use this service.

The protocol makes no assumptions as to the contents of the SONET/SDH SPE, and therefore is applicable to SONET/SDH circuits carrying any type of payload.

Because the protocol terminates the SONET/SDH section and line before emulating the individual SPEs, the protocol allows the PSN to operate as a distributed SONET/SDH cross-connect.

3.1 Fidelity of Emulated SONET/SDH SPE services

The protocol does not make any assumptions about the capabilities of the underlying PSN. However, the fidelity of the emulated service will be dependent on the characteristics of the underlying PSN.

Emulated SONET/SDH SPE services may differ from native SONET/SDH services on the following parameters: SPE timing, service reliability, end-to-end delay, and bit-error-rate. Each of these parameters is discussed below.

Because of the rigorous synchronization requirements implied by SONET/SDH services, it is expected that the protocol will most commonly be deployed in situations where a common timing reference is available at the PW end-points. Large network operators have well-defined methods for distributing Stratum timing references (such as BITS, SASE, or GPS). Using these references is the most direct technique that can be mathematically proven to meet the relevant network synchronization specifications.

However, smaller network operators or remote locations in larger networks may not have access to a common reference either by design or due to a persistent fault in the timing distribution network. In the absence of common references adaptive timing recovery techniques may be employed. However, the fidelity of the recovered SPE timing will be dependent on the packet-delay variation behavior of the underlying PSN and the robustness of the timing recovery algorithm used. As a result, it may be difficult in these circumstances to mathematically prove that the recovered SPE timing is in compliance with relevant synchronization standards.

Service Reliability may be impacted by two components: the robustness of the underlying PSN and whether specific steps have been taken to protect the emulated service (such as 1+1 protection switching on the emulated service). The jitter buffer and packet reordering mechanisms associated with the protocol increase resilience of the emulated service to fast PSN rerouting events.

End-to-end delay will be impacted by both the transit delay through the PSN and the packet-delay-variation characteristics of the PSN. The protocol makes no assumption regarding either of these parameters. However, the tighter the bound on transit delay and delay variation, the shorter the end-to-end delay of the emulated circuit will be.

BER for emulated circuits will be dependent on the characteristics of the PSN. Each packet dropped by the PSN will result in an equivalent number of byte errors on the emulated SPE. Using smaller packet sizes can reduce the effect of lost packets on the emulated service but increases the ratio of overhead to payload. The protocol allows flexibility in packet length to accommodate the desired BER/Overhead working point.

To the extent possible, the use of low-loss paths (for example, by reserving link bandwidth and router/switch buffering) in the PSN will enhance the fidelity of the emulated circuits.

3.2 Performance Monitoring and Fault Isolation

The protocol allows collection of SONET/SDH-like faults and performance monitoring parameters. Similarity with existing SONET/SDH services is increased by the protocol's ability to carry 'far end error' indications (i.e. RDI). The protocol performance monitoring capabilities are based on SONET/SDH requirements as reflected by the available standards, and adapted to the nature of the protocol.

The protocol provides the ability to detect lost packets and hence allows it to distinguish between PSN problems and problems external to the PSN as causes of outages and/or degradations of the emulated service. In addition, the protocol supports fast detection of defects, enabling vendors to implement rapid fault recovery mechanisms for the emulated circuit.

3.3 Other Considerations

The protocol allows for bandwidth conservation in the PSN by carrying only AIS-P and/or STS SPE Unequipped indications instead of empty payloads, thus providing for efficiency gains on the PW. Additional payload conservation techniques may be defined in the future.

Being a constant bit rate (CBR) service, the protocol cannot provide TCP-friendly behavior under network congestion. It will operate best in environments where the Diff-Serv EF PHB with allocated bandwidth is available end-to-end between the PW endpoints and the EF bandwidth is sized to meet the requirements of the emulated SONET/SDH circuits, or over a well engineered path as available through the relevant signaling protocols like RSVP-TE and CR-LDP for MPLS PSNs. Using these methods will prevent contention between the SONET Emulation protocol and TCP traffic. Unusable service characteristics from the packet switched network may be used to trigger circuit/PW teardown or switch-over.

4 Scope

This document describes how to provide CEP for the following digital signals:

1. SONET STS-1 synchronous payload envelope (SPE)/SDH VC-3

2. STS-Nc SPE (N = 3, 12, 48, or 192)/SDH VC-4, VC-4-4c, VC-4-16c, or VC-4-64c

For the remainder of this document, these constructs will be referred to as SONET/SDH channels.

Although this document currently covers up to OC-192c/VC-4-64c, future revision MAY address higher rates.

Other SONET/SDH signals, such as virtual tributary (VT) structured sub-rate mapping, are not explicitly discussed in this document; however, it can be extended in the future to support VT and lower speed non-SONET/SDH services.

5 CEP Encapsulation Format

In order to transport SONET/SDH SPEs through a packet-oriented network, the SPE is broken into fragments. A CEP Header is pre-pended to each fragment. The resulting packet is encapsulated in RTP for transmission over an arbitrary PSN.

(Note: under certain circumstances the RTP header may be suppressed to conserve network bandwidth. See section 5.4.3 for details).

The basic CEP packet appears in Figure 1.

```
+-----------------------------------+
|   PSN and Multiplexing Layer      |
|             Headers               |
+-----------------------------------+
|           RTP Header              |
|           (RFC1889)               |
+-----------------------------------+
|           CEP Header              |
+-----------------------------------+
|                                   |
|                                   |
|        SONET/SDH SPE Fragment     |
|                                   |
|                                   |
+-----------------------------------+
```

Figure 1 - Basic CEP Packet

5.1 SONET/SDH SPE Fragment

The SONET/SDH Fragments MUST be byte aligned with the SONET/SDH SPE.

The first bit received from each byte of the SONET/SDH SPE MUST be the Most Significant Bit of each byte in the SONET/SDH SPE fragment.

SONET/SDH bytes are placed into the SONET/SDH fragment in the same order in which they are received.

SONET/SDH optical interfaces use binary coding and therefore are scrambled prior to transmission to insure an adequate number of transitions. For clarity, this scrambling will be referred to as physical layer scrambling/descrambling.

In addition, many payload formats (such as for ATM and HDLC) include an additional layer of scrambling to provide protection against transition density violations within the SPEs. This function will be referred to as payload scrambling/descrambling.

CEP assumes that physical layer scrambling/descrambling occurs as part of the SONET/SDH section/line termination Native Service Processing (NSP) functions.

However, CEP makes no assumption about payload scrambling. The SONET/SDH SPE fragments MUST be constructed without knowledge or processing of any incidental payload scrambling.

5.2 CEP Header

The CEP Header supports a basic and extended mode. The Basic CEP Header provides the minimum functionality necessary to accurately emulate a TDM SONET over a PSN if a common reference is available at both ends of the PW.

Enhanced functionality and commonality with other real-time Internet applications is provided by RTP encapsulation.

Bit 0 of the first 32-bit CEP header indicates whether or not the extended header is present. When this bit is 0, then no extended header is present. When this bit is 1, then an extended header is present. At this time, the contents of the extended header are for future study. However, it is expected that this field will provide support for payload compression, header protection, enhanced performance monitoring, and/or other extensions to the base protocol.

The Basic CEP header has the following format:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |0|R|D|N|P| Structure Pointer[0:12] |  Sequence Number[0:13]    |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 2 - Basic CEP Header Format

The Extended CEP header appears below:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |1|R|D|N|P| Structure Pointer[0:12] |  Sequence Number[0:13]    |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |                          Reserved                               |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 3 - Extended CEP Header Format

The above fields are defined as follows:

R bit: CEP-RDI. This bit is set to one to signal to the remote CEP function that a loss of packet synchronization has occurred. See section 6.4 for details.

D bit: Signals DBA Mode. MUST be set to zero for Normal Operation. MUST be set to one if CEP is currently in DBA mode. DBA is an optional mode during which trivial SPEs are not transmitted into the packet network. See Table 1 and section 6 for further details.

The N and P bits: MAY be used to explicitly relay negative and positive pointer adjustment events across the PSN. They are also used to relay SONET/SDH maintenance signals such as AIS-P. See Table 1 and sections 7 and 9 for more details.

D	N	P	Interpretation
0	0	0	Normal Mode - No Ptr Adjustment
0	0	1	Normal Mode - Positive Ptr Adjustment
0	1	0	Normal Mode - Negative Ptr Adjustment
0	1	1	Normal Mode - AIS-P
1	0	0	DBA Mode - STS SPE Unequipped
1	0	1	DBA Mode - STS SPE Unequipped Pos Ptr Adj
1	1	0	DBA Mode - STS SPE Unequipped Neg Ptr Adj
1	1	1	DBA Mode - AIS-P

Table 1. Interpretation of D, N, and P bits

Sequence Number[0:13]: This is a packet sequence number, which MUST continuously cycle from 0 to 0x3FFF. It is generated and processed in accordance with the rules established in [RFC1889]. When the RTP header is used, this sequence number MUST match the LSBs of the RTP sequence Number.

Structure Pointer[0:12]: The Structure Pointer MUST contain the offset of the J1 byte within the CEP SPE Fragment. The value is from 0 to 0x1FFE, where 0 means the first byte after the CEP header. The Structure Pointer MUST be set to 0x1FFF if a packet does not carry the J1 byte. See [SONET], [GR253], and [G707] for more information on the J1 byte and the SONET/SDH payload pointer. Implementations MUST support SPE Fragments of 783 bytes and MAY support SPE fragments of from 8 to 8191 bytes.

Note 1: Implementations that choose to support programmable payload lengths SHOULD support payloads that are an integer multiple of 8 bytes.

Note 2: CEP packets are fixed in length for all of the packets of a particular emulated TDM stream. This length is statically provisioned for each TDM stream. Therefore, the length of each CEP packet does not need to be carried in the CEP header.

5.3 RTP Header

CEP uses the fixed RTP Header as shown below.

```
    0                   1                   2                   3
    0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |V=2|P|X|  CC   |M|     PT      |       sequence number         |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |                           timestamp                           |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |           synchronization source (SSRC) identifier            |
   +=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+
```

o V (version) is always set to 2

o P (padding) is always set to 0

o X (header extension) is always set to 0

o CC (CSRC count) is always set to 0

o M (marker) is set to 0 for CEP packets.

o PT (payload type) is used to identify packets carrying the packetized SONET/SDH data. One PT value should be allocated from the range of dynamic values (see [RTP-TYPES]) for every CEP PW. Allocation is done during the PW setup and MUST be the same for both PW directions. The PE at the PW ingress MUST set the PT value in the RTP header to the allocated value.

o Sequence Number is used primarily to provide the common PW sequencing function as well as detection of lost packets. It is generated and processed in accordance with the rules established in [RFC1889].

o Timestamp is used primarily for carrying timing information over the network. Their values are used in accordance with the rules established in [RFC1889]. Frequency of the clock used for generating timestamps MUST be 19.44 MHz based on a local reference.

O SSRC (synchronization source) value in the RTP header MAY be used for detection of misconnections.

5.4 PSN Encapsulation

In principle, CEP packets can be carried over any packet-oriented network. The following sections describe specifically how CEP packets MUST be encapsulated for carriage over MPLS or IP networks.

5.4.1 IP Encapsulation

CEP uses the standard IP/UDP/RTP encapsulation scheme as shown below. The UDP destination port MUST be used to Demultiplex individual SONET channels.

```
+-----------------------------------+
|                                   |
|          IPv6/v4 Header           |
|                                   |
+-----------------------------------+
|             UDP Header            |
+-----------------------------------+
|             RTP Header            |
+-----------------------------------+
|             CEP Header            |
+-----------------------------------+
|                                   |
|                                   |
|       SONET/SDH SPE Fragment      |
|                                   |
|                                   |
+-----------------------------------+
```

Figure 4 - IP Transport Encapsulation

5.4.2 MPLS Encapsulation

RTP MAY be directly encapsulated in MPLS as shown below. To transport a CEP packet over an MPLS network, an MPLS label-stack MUST be pushed on top of the CEP packet. The bottom label in the MPLS label stack MUST be used to demultiplex individual SONET channels. In keeping with the conventions used in [MARTINI-TRANS], this demultiplexing label is referred to as the VC Label and the upper labels are referred to as Tunnel Labels.

```
+-----------------------------------+
|  One or more MPLS Tunnel Labels   |
+-----------------------------------+
|            VC Label               |
+-----------------------------------+
|           RTP Header              |
+-----------------------------------+
|           CEP Header              |
+-----------------------------------+
|                                   |
|                                   |
|       SONET/SDH SPE Fragment      |
|                                   |
|                                   |
+-----------------------------------+
```

Figure 5 - Typical MPLS Transport Encapsulation

5.4.3 RTP Header Suppression

In addition to normal RTP header compression mechanisms as described in [RFC2508] and [RFC3095], an additional option may be used in CEP which suppresses transmission of the RTP header altogether.

This mode may be used when both SONET Emulation PEs have access to a common reference clock and both support RTP Header Suppression. Under these conditions the following encapsulation formats may be used.

The choice to utilize RTP Header Suppression may be statically configured using [CEM-MIB], or signaled using a PW maintenance protocol such as [MARTINI-TRANS].

```
              +-----------------------------------+
              |                                   |
              |         IPv6/v4 Header            |
              |                                   |
              +-----------------------------------+
              |            UDP Header             |
              +-----------------------------------+
              |            CEP Header             |
              +-----------------------------------+
              |                                   |
              |                                   |
              |       SONET/SDH SPE Fragment      |
              |                                   |
              |                                   |
              +-----------------------------------+
```

Figure 6 - IP Transport Encapsulation w/ RTP Header Suppression

```
                 +-----------------------------------+
                 |  One or more MPLS Tunnel Labels   |
                 +-----------------------------------+
                 |            VC Label               |
                 +-----------------------------------+
                 |           CEP Header              |
                 +-----------------------------------+
                 |                                   |
                 |                                   |
                 |       SONET/SDH SPE Fragment      |
                 |                                   |
                 |                                   |
                 +-----------------------------------+
```

Figure 7 - MPLS Transport Encapsulation w/ RTP Header Suppression

5.5 L2TP Encapsulation

Encapsulation for L2TP PSNs is for future study.

6 CEP Operation

The following sections describe CEP operation.

6.1 Introduction and Terminology

CEP MUST support a normal mode of operation and MAY support an optional extension called Dynamic Bandwidth Allocation (DBA). During normal operation, SONET/SDH payloads are fragmented, pre-pended with the appropriate headers and then transmitted into the packet network. During DBA mode, only the headers are transmitted. This is done to conserve bandwidth when meaningful user data is not present in the SPE, such as during AIS-P or STS SPE Unequipped.

6.1.1 CEP Packetizer and De-Packetizer

As with all adaptation functions, CEP has two distinct components: adapting TDM SONET/SDH into a CEP packet stream, and converting the CEP packet stream back into a TDM SONET/SDH. The first function will be referred to as CEP Packetizer and the second as CEP De-Packetizer. This terminology is illustrated in Figure 8.

```
          +------------+            +---------------+
          |            |            |               |
SONET --> |    CEP     | --> PSN  --> |      CEP      | --> SONET
 SDH      | Packetizer |            | De-Packetizer |      SDH
          |            |            |               |
          +------------+            +---------------+
```

Figure 8 - CEP Terminology

Note: the CEP de-packetizer requires a buffering mechanism to account for delay variation in the CEP packet stream. This buffering mechanism will be generically referred to as the CEP jitter buffer.

6.1.2 CEP DBA

DBA is an optional mode of operation that only transmits the headers into the packet network under certain circumstances such as AIS-P or STS Unequipped.

If DBA is supported by a CEP implementation, the user SHOULD be able to configure if DBA will be triggered by AIS-P, STS Unequipped, both, or neither on a per channel basis.

If DBA is supported, the determination of AIS-P and STS Unequipped MUST be based on the state of SONET/SDH Section, Line, and Path Overhead bytes.

During AIS-P, there is no valid payload pointer, so pointer adjustments cannot occur. During STS Unequipped, the SONET/SDH payload pointer is valid, and therefore pointer adjustments MUST be supported even during DBA. See Table 1 for details.

6.2 Description of Normal CEP Operation

During normal operation, the CEP packetizer will receive a fixed rate byte stream from a SONET/SDH interface. When a packets worth of data has been received from a SONET/SDH channel, the necessary headers are pre-pended to the SPE fragment and the resulting CEP packet is transmitted into the packet network. Because all CEP packets associated with a specific SONET/SDH channel will have the same length, the transmission of CEP packets for that channel SHOULD occur at regular intervals.

At the far end of the packet network, the CEP de-packetizer will receive packets into a jitter buffer and then play out the received byte stream at a fixed rate onto the corresponding SONET/SDH channel. The jitter buffer SHOULD be adjustable in length to account for varying network delay behavior. The receive packet rate from the packet network should be exactly balanced by the transmission rate onto the SONET/SDH channel, on average. The time over which this average is taken corresponds to the depth of the jitter buffer for a specific CEP channel.

The RTP sequence numbers provide a mechanism to detect lost and/or mis-ordered packets. The sequence number in the CEP header may be used when transmission of the RTP header is suppressed (see section 5.4.3 for details). The CEP de-packetizer MUST detect lost or mis-ordered packets. The CEP de-packetizer SHOULD play out an all ones pattern (AIS) in place of any dropped packets. The CEP de-packetizer MAY re-order packets received out of order. If the CEP de-packetizer does not support re-ordering, it must drop mis-ordered packets.

6.3 Description of CEP Operation during DBA

There are several issues that should be addressed by a workable CEP DBA mechanism. First, when DBA is invoked, there should be a substantial savings in bandwidth utilization in the packet network. The second issue is that the transition in and out of DBA should be tightly coordinated between the local CEP packetizer and CEP de-packetizer at the far side of the packet network. A third is that the transition in and out of DBA should be accomplished with minimal disruption to the adapted data stream.

Another goal is that the reduction of CEP traffic due to DBA should not be mistaken for a fault in the packet network or vice-versa. Finally, the implementation of DBA should require minimal modifications beyond what is necessary for the nominal CEP case. The mechanism described below is a reasonable balance of these goals.

During DBA, packets MUST be emitted at exactly the same rate as they would be during normal operation. This SHOULD be accomplished by transmitting each DBA packet after a complete packet of data has been received from the SONET/SDH channel. The only change from normal operation is that the CEP packets during DBA MUST only suppress the transmission of the SPE while still sending the appropriate headers. Because some links have a minimum supported packet size, the CEP packetizer MAY append a configurable number of bytes immediately after the CEP header to pad out the CEP packet to reach the mimumum supported packet size. The D-bit MUST be set to one, to indicate that DBA is active.

The CEP de-packetizer MUST assume that each packet received with the D-bit set represents a normal-sized packet containing an AIS-P or SPE Unequipped payload as noted by N and P. See Table 1. The CEP de-packetizer MUST accept DBA packets with or without padding.

This allows the CEP packetization and de-packetization logic during DBA to be similar to the nominal case. It ensures that the correct SONET/SDH indication is reliably transmitted between CEP adaptation points. It minimizes the risk of under or over running the jitter buffer during the transition in and out of DBA, since packets are continuously transmitted during DBA. And, it guarantees that faults in the packet network are recognized as distinctly different from line conditioning on the SONET/SDH interfaces.

6.4 Packet Synchronization

A key component in declaring the state of a CEP service is whether or not the CEP de-packetizer is in or out of packet synchronization. The following paragraphs describe how that determination is made.

As discussed in section 6, a CEP de-packetizer MAY or MAY NOT support re-ordering of mis-ordered packets.

As packets are received from the PSN, they are placed into a jitter buffer prior to play out on the SONET interface. If a CEP de-packetizer supports re-ordering, any packet received before its play out time will still be considered valid.

If a CEP de-packetizer does not support re-ordering, a number of approaches may be used to minimize the impact of mis-ordered or lost packets on the final re-assembled SONET stream. For example, [AAL1] uses a simple state-machine to re-order packets in a sub-set of possible cases.

However, the final determination as to whether or not to declare acquisition or loss of packet synchronization MUST be based on the same criteria regardless of whether an implementation supports or does not support re-ordering.

Therefore, the determination of acquisition or loss of packet synchronization is always made at SONET play-out time. During SONET play-out, the CEP de-packetizer will play received CEP packets onto the SONET interface. However, if the jitter buffer is empty or the packet to be played out has not been received, the CEP de-packetizer will play out an empty packet onto the SONET interface in place of the unavailable packet.

The acquisition of packet synch is based on the number of sequential CEP packets that are played onto the SONET interface. While, loss of packet synch is based on the number of sequential 'empty' packets that are played onto the SONET interface. Specific details of these two cases is described below.

6.4.1 Acquisition of Packet Synchronization

At startup, a CEP de-packetizer will be out of packet synchronization by default. To declare packet synchronization at startup or after a loss of packet synchronization, the CEP de-packetizer must play-out a configurable number of CEP packets with sequential sequence numbers towards the SONET interface.

6.4.2 Loss of Packet Synchronization

Once a CEP de-packetizer is in packet sync, it may encounter a set of events that will cause it to lose packet synchronization.

If the CEP de-packetizer encounters more than a configurable number of sequential empty packets, the CEP de-packetizer MUST declare loss of packet synchronization (LOPS) defect.

Loss of Packet Synchronization (LOPS) failure is declared after 2.5 +/- 0.5 seconds of LOPS defect, and cleared after 10 seconds free of LOPS defect state. The VC is considered down as long as LOPS failure is declared.

7 SONET/SDH Maintenance Signals

There are several issues that must be considered in the mapping of maintenance signals between SONET/SDH and a PSN. A description of how these signals and conditions are mapped between the two domains is described below.

For clarity, the mappings are split into two groups: SONET/SDH to PSN, and PSN to SONET/SDH.

7.1 SONET/SDH to PSN

The following sections describe how SONET/SDH Maintenance Signals and Alarm conditions are mapped into a Packet Switched Network.

7.1.1   AIS-P Indication

In a SONET/SDH network, SONET Path outages are signaled using maintenance alarms such as Path AIS (AIS-P). In particular, AIS-P indicates that the SONET/SDH Path is not currently transmitting valid end-user data, and the SPE contains all ones.

It should be noted that nearly every type of service-affecting section or line defect will result in an AIS-P condition.

The SONET/SDH hierarchy is illustrated below.

```
                               +----------+
                               |   PATH   |
                               +----------+
                                    ^
                                    |
                                  AIS-P
                                    |
                                    |
                               +----------+
                               |   LINE   |
                               + ---------+
                                  ^     ^
                                  |     |
                                AIS-L   +------ LOP
                                  |
                                  |
                               +----------+
                               | SECTION  |
                               +----------+
                                  ^    ^
                                  |    |
                                  |    |
                                 LOS  LOF
```

Figure 9 - SONET/SDH Fault Hierarchy

Should the Section Layer detect a Loss of Signal (LOS) or Loss of Frame (LOF) condition, it sends AIS-L up to the Line Layer. If the Line Layer detects AIS-L or Loss of Path (LOP), it sends AIS-P to the Path Layer.

In normal mode during AIS-P, CEP packets are generated as usual. The N and P bits MUST be set to 11 binary to signal AIS-P explicitly through the packet network. The D-bit MUST be set to zero to indicate that the SPE is being carried through the packet network.

Normal CEP packets with the SPE fragment, CEP Header, the Circuit ID Word, and PSN Header MUST be transmitted into the packet network.

However, to conserve network bandwidth during AIS-P, DBA MAY be employed. If DBA has been enabled for AIS-P and AIS-P is currently occurring, the N and P bits MUST be set to 11 binary to signal AIS, and the D-bit MUST be set to one to indicate that the SPE is not being carried through the packet network. Only the CEP header, the Circuit ID Word, and the PSN Header MUST be transmitted into the packet network.

7.1.2 STS SPE Unequipped Indication

The declaration of STS SPE unequipped MUST conform to [GR253]. Quoted below:

"R6-135 [481] STS PTE shall detect an STS Path Unequipped (UNEQ-P) defect within 10 ms of the onset of at least five consecutive samples (which may or may not be consecutive frames) of unequipped STS Signal Labels (C2 byte), as specified in Table 6-2"

The termination of STS SPE unequipped MUST also conform to [GR253].

"R6-137 [485v2] STS PTE shall terminate an UNEQ-P defect within 10 ms of the onset of at least five consecutive samples (which may or may not be consecutive frames) of STS Signal Labels that are not unequipped or all-ones, as specified in Table 6-2"

For normal operation during SPE Unequipped, the N and P bits MUST be interpreted as usual. The SPE MUST be transmitted into the packet network along with the appropriate headers, and the D-Bit MUST be set to zero.

If DBA has been enabled for STS SPE Unequipped and the Unequipped is occurring on the SONET/SDH channel, the D-bit MUST be set to one to indicate DBA is active. Only the necessary headers are transmitted into the packet network. The N and P bits MAY be used to signal pointer adjustments as normal. See Table 1 and section 6 for details.

7.1.3 CEP-RDI

The CEP function MUST send CEP-RDI towards the packet network during loss of packet synchronization. This MUST be accomplished by setting the R bit to one in the CEP header.

7.2 PSN to SONET/SDH

The following sections discuss how the various conditions on the packet network are converted into SONET/SDH indications.

7.2.1 AIS-P Indication

There are several conditions in the packet network that will cause the CEP de-packetization function to play out an AIS-P indication towards a SONET/SDH channel.

The first of these is the receipt of CEP packets with the N and P bits set to one, and the D-bit set to zero. This is an explicit indication of AIS-P being received at the far-end of the packet network, with DBA disabled for AIS-P. The CEP de-packetizer MUST play out the received SPE fragment (which will incidentally be carrying all ones), and MUST configure the SONET/SDH Overhead to signal AIS-P as defined in [SONET], [GR253], and [G707].

The second case is the receipt of CEP packets with the N and P bits set to one, and the D-bit set to one. This indicates that AIS-P is being received at the far-end of the packet network, with DBA enabled for AIS-P. The CEP de-packetizer MUST play out one packet's worth of all ones for each packet received, and MUST configure the SONET/SDH Overhead to signal AIS-P as defined in [SONET], [GR253], and [G707].

A third case that will cause a CEP de-packetization function to play out an AIS-P indication onto a SONET/SDH channel is during loss of packet synchronization. The CEP de-packetizer MUST configure the SONET/SDH Overhead to signal AIS-P as defined in [SONET], [GR253], and [G707].

7.2.2 STS SPE Unequipped Indication

There are three conditions in the packet network that will cause the CEP function to transmit STS SPE Unequipped indications onto the SONET/SDH channel.

The first, which is transparent to CEP, is the receipt of regular CEP packets that happen to be carrying an SPE that contains the appropriate Path overhead to signal STS SPE unequipped. This case does not require any special processing on the part of the CEP de-packetizer.

The second case is the receipt of CEP packets that have the D-bit set to one to indicate DBA active and the N and P bits set to 00 binary, 01 binary, or 10 binary to indicate SPE Unequipped with or without pointer adjustments. The CEP de-packetizer MUST use this information to transmit a packet of all zeros onto the SONET/SDH interface, and adjust the payload pointer as necessary.

The third case when a CEP de-packetizer MUST play out an STS SPE Unequipped Indication towards the SONET interface is when the VC-label has been withdrawn due to de-provisioning of the circuit.

8 SONET/SDH Transport Timing

It is assumed that the distribution of SONET/SDH Transport timing information is addressed through external mechanisms such as Building Integrated Timing System (BITS), Stand Alone Synchronization Equipment (SASE), Global Positioning System (GPS) or other such methods and is therefore outside of the scope of this specification.

9 SONET/SDH Pointer Management

A pointer management system is defined as part of the definition of SONET/SDH. Details on SONET/SDH pointer management can be found in [SONET], [GR253], and [G707]. If there is a frequency offset between the frame rate of the transport overhead and that of the SONET/SDH SPE, then the alignment of the SPE shall periodically slip back or advance in time through positive or negative stuffing.

The emulation of this aspect of SONET networks may be accomplished using a variety of techniques including (but not limited to) explicit pointer adjustment relay (EPAR) and adaptive pointer management (APM).

In any case, the handling of the SPE data by the CEP packetizer is the same.

During a negative pointer adjustment event, the CEP packetizer MUST incorporate the H3 byte from the SONET/SDH stream into the CEP packet payload in order with the rest of the SPE. During a positive pointer adjustment event, the CEP de-packetizer MUST strip the stuff byte from the CEP packet payload.

When playing out a negative pointer adjustment event, the appropriate byte of the CEP payload MUST be placed into the H3 byte of the SONET/SDH stream. When playing out a positive pointer adjustment, the CEP de-packetizer MUST insert a stuff-byte into the appropriate position within the SONET/SDH stream.

The details regarding the use of the H3 byte and stuff byte during positive and negative pointer adjustments can be found in [SONET], [GR253], and [G707].

9.1 Explicit Pointer Adjustment Relay (EPAR)

CEP provides an OPTIONAL mechanism to explicitly relay pointer adjustment events from one side of the PSN to the other. This technique will be referred to as Explicit Pointer Adjustment Relay (EPAR). The mechanics of EPAR are described below.

The following text only applies to implementations that choose to implement EPAR. Any CEP implementation that does not support EPAR

MUST either set the N and P bits to zero or utilize them to relay AIS-P and STS Unequipped as shown in table 1.

If EPAR is being used, the pointer adjustment event MUST be transmitted in three consecutive packets by the packetizer. The de-packetizer MUST play out the pointer adjustment event when any one packet with N/P bit set is received.

References [SONET], [GR253], and [G707] specify that pointer adjustment events MUST be separated by three SONET/SDH frames without a pointer adjustment event. In order to explicitly relay all legal pointer adjustment events, the packet size for a specific circuit SHOULD be no larger than (783 * 4 * N)/3, where N is the STS-Nc multiplier.

However, there are SONET implementations that allow pointer adjustments to occur in back to back SONET/SDH frames. In order to support this possibility, EPAR implementations SHOULD set the packet size for a particular circuit to be no larger than (783*N)/3. Where N is the STS-Nc multiplier.

Since the minimum value of N is one, EPAR implementations SHOULD support a minimum payload length of 783/3 or 261 bytes.

For EPAR implementations, the CEP de-packetizer MUST utilize the CEP sequence numbers to insure that SONET/SDH pointer adjustment events are not played any more frequently than once per every three CEP packets transmitted by the remote CEP packetizer.

If both bits are set, then an AIS-P event has occurred (this is further discussed in section 7).

When DBA is invoked (i.e. the D-bit = 1), N and P have additional meanings. See Table 1 and section 6.

9.2 Adaptive Pointer Management (APM)

Another OPTIONAL method that may be used to emulate SONET pointer management is Adaptive Pointer Management (APM). In basic terms, APM uses information about the depth of the CEP jitter buffers to introduce pointer adjustments in the reassembled SONET SPE.

Details about specific APM algorithms is for future study.

10 CEP Performance Monitors

SONET/SDH as defined in [SONET], [GR253], and [G707] includes the definition of several counters that may be used to monitor the performance of SONET/SDH services. These counters are referred to as Performance Monitors.

In order for CEP to be utilized by traditional SONET/SDH network operators, CEP SHOULD provide similar functionality. To this end, the following sections describe a number of counters that will collectively be referred to as CEP Performance Monitors.

10.1   Near-End Performance Monitors

These performance monitors are maintained by the CEP De-Packetizer during reassembly of the SONET stream.

The performance monitors are based on two types of defects.

Type 1 defect is defined as: missing or dropped packet.
Type 2 defect is defined as: buffer under run, buffer over-run, LOPS.

The specific performance monitors that are defined for CEP are as follows:

```
ES-CEP        - CEP Errored Seconds
SES-CEP       - CEP Severely Errored Seconds
UAS-CEP       - CEP Unavailable Seconds
```

Each second that contain at least one type 1 defect SHALL be declared as ES-CEP.

Each second that contain type 2 defect, or missing packets above pre-defined, configurable threshold of missing/dropped packets SHALL be declared both SES-CEP and ES-CEP. Default value for missing packet to SES is 3.

UAS-CEP SHALL be declared after X consecutives SES-CEP, cleared after X consecutive seconds without SES-CEP. Default value of X is 10 seconds.

Once unavailability is declared, ES and SES counts SHALL be inhibited up to the point where the unavailability was started. Once unavailability is removed, ES that occurred along the X seconds clearing period SHALL be added to the ES counts. An update is required even for closed intervals if necessary.

FC-CEP is the number of time type 1 or type 2 defect states were declared. The NE SHALL have thresholding on ES-CEP, SES-CEP and UAS-CEP (thresholding mean activate a notification if more than pre-defined # of seconds are declared as ES, etc. in 15 minutes interval).

10.2 Far-End Performance Monitors

These performance monitors provide insight into the CEM De-packetizer at the far-end of the PSN.

Far end statistics are based on the RDI-CEP bit. Limited functionality is supported compared to [GR-253] for simplicity and because it is assumed that all relevant statistics are available from the end point of the PW. CEP-FE defect is declared when CEP-RDI is set in the incoming CEP packets.

CEP-FE failure declared after 2.5 +/- 0.5 seconds of CEP-FE defect, and cleared after 10 seconds free of CES-FE defect state. Sending notification to the OS for CEP-FE failure is local policy.

This draft does not attempt to define SES-CEPFE, UAS-CEPFE and FC-CEPFE, but they can be added if to fully emulate GR-253 far end PM (thresholding is required too here except for FC-CEPFE). (Note that ES-CEPFE is not relevant since CEP does not report back missing packets - only LOPS which is SES).

The definition of additional performance monitors is for future study.

11 Open Issues

This version of the draft does not address payload compression within the emulated SONET. Payload compression is expected to be supported by future versions of this draft by utilizing the extended CEP header.

This version of the draft does not tie into PWE3 maintenance mechanisms for the setup and tear down of services. That short-coming will be addressed in future revisions of this document.

Underlying MPLS QoS requirements are not covered by this revision of the draft. Future revisions may discuss underlying QoS requirements.

Support for VT and lower speed non-SONET/SDH services are not covered in this revision of the draft. Future revisions may address VT and non-SONET/SDH TDM services.

An alternate version of DBA has been suggested that would suppress transmission of the entire CEP packet stream under certain circumstances. Future versions of this draft may define such a mechanism.

It is possible to define SONET Emulation specific redundancy mechanisms, such as 1+1 or N:1. Future versions of this draft may define such mechanisms.

12 Security Considerations

This document does not address or modify security issues within the relevant PSNs.

13 Intellectual Property Disclaimer

This document is being submitted for use in IETF standards discussions. Vivace Networks, Inc. has filed one or more patent applications relating to the CEP technology outlined in this document. Vivace Networks, Inc. will grant free unlimited licenses for use of this technology.

14 References

[RFC2026] Bradner, S., "The Internet Standards Process -- Revision 3", BCP 9, RFC2026, October 1996.

[PWE3-REQ] XiPeng Xiao et al, Requirements for Pseudo Wire Emulation Edge-to-Edge (PWE3), Work in Progress, July-2001, draft-ietf-pwe3-requirements-01.txt

[PWE3-FW] Prayson Pate et al, Framework for Pseudo Wire Emulation Edge-to-Edge (PWE3), Work in progress, February 2002, draft-ietf-pwe3-framework-00.txt

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[PWE3-LAYERS], Stewart Bryant et al., Protocol Layering in PWE3, Work in Progress, February 2002, pwe3-protocol-layering-01.txt

[SONET] American National Standards Institute, "Synchronous Optical Network (SONET) - Basic Description including Multiplex Structure, Rates and Formats," ANSI T1.105-1995.

[GR253] Telcordia Technologies, "Synchronous Optical Network (SONET) Transport Systems: Common Generic Criteria", GR-253-CORE, Issue 3, September 2000.

[G707] ITU Recommendation G.707, "Network Node Interface For The Synchronous Digital Hierarchy", 1996.

[RFC1889] H. Schulzrinne et al, RTP: A Transport Protocol for Real-Time Applications, RFC 1889, IETF, 1996

[ROHC-LLA] Lars-Eric Jonsson et al, A Link-Layer Assisted ROHC Profile for IP/UDP/RTP draft-ietf-rohc-rtp-lla-03.txt.

[CEM] Malis et al, "SONET/SDH Circuit Emulation Service Over MPLS (CEM) Encapsulation", draft-malis-sonet-ces-mpls-05.txt, work in progress, July 2001.

[CEM-MIB] Danenberg et al, "SONET/SDH Circuit Emulation Service Over PSN (CEP) Management Information Base Using SMIv2", draft-danenberg-pw-cem-mib-02.txt, work in progress, Feb 2002.

[MARTINI-TRANS] Martini et al, "Transport of Layer 2 Frames Over MPLS", draft-martini-l2circuit-trans-mpls-06.txt, work in progress, July 2001.

[MARTINI-ENCAP] Martini et al, "Encapsulation Methods for Transport of Layer 2 Frames Over MPLS", draft-martini-l2circuit-encap-mpls-02.txt, work in progress, July 2001.

[CESoPSN] Vainshtein et al, "TDM Circuit Emulation Service over Packet Switched Network", draft-vainshtein-cesopsn-02.txt, work in progress, February 2002.

[CES-VT] Pate et al, "TDM Service Specification for Pseudo-Wire Emulation Edge-to-Edge", draft-pate-pwe3-tdm-03.txt, work in progress, January 2001.

[RFC2508] S.Casner, V.Jacobson, Compressing IP/UDP/RTP Headers for Low-Speed Serial Links, RFC 2508, IETF, 1999

[RFC3095] C.Bormann (Ed.), RObust Header Compression (ROHC): Framework and four profiles: RTP, UDP, ESP, and uncompressed, RFC 3095, IETF, 2001

[AAL1] ITU-T, "Recommendation I.363.1, B-ISDN Adaptation Layer Specification: Type AAL1", Appendix III, August 1996.

15 Acknowledgments

The authors would like to thank all of the members of the PWE3 working group who have contributed to the development of this draft, and specifically Danny McPherson and Allison Mankin for their advice and assistance.

16 Author's Addresses

Andrew G. Malis
Vivace Networks, Inc.
2730 Orchard Parkway
San Jose, CA 95134
Email: Andy.Malis@vivacenetworks.com

Ken Hsu
Vivace Networks, Inc.
2730 Orchard Parkway
San Jose, CA 95134
Email: Ken.Hsu@vivacenetworks.com

Jeremy Brayley
Laurel Networks, Inc.
2706 Nicholson Rd.
Sewickley, PA 15143
Email: jbrayley@laurelnetworks.com

Steve Vogelsang

Laurel Networks, Inc.
2706 Nicholson Rd.
Sewickley, PA 15143

Email: sjv@laurelnetworks.com

John Shirron
Laurel Networks, Inc.
2607 Nicholson Rd.
Sewickley, PA 15143
Email: jshirron@laurelnetworks.com

Luca Martini
Level 3 Communications, LLC.
1025 Eldorado Blvd.
Broomfield, CO 80021
Email: luca@level3.net

Tom Johnson
Litchfield Communications, Inc.
76 Westbury Park Rd.
Watertown, CT 06795
Email: tom_johnson@litchfieldcomm.com

Ed Hallman
Litchfield Communications, Inc.
76 Westbury Park Rd.
Watertown, CT 06795
Email: ed_hallman@litchfieldcomm.com

Marlene Drost
Litchfield Communications, Inc.
76 Westbury Park Rd.
Watertown, CT 06795
Email: marlene_drost@litchfieldcomm.com

Jim Boyle
Protocol Driven Networks, Inc.
1381 Kildaire Farm #288
Cary, NC 27511
Email: jboyle@pdnets.com

David Zelig
Corrigent Systems LTD.
126, Yigal Alon st.
Tel Aviv, ISRAEL
Email: davidz@corrigent.com

Ron Cohen
Lycium Networks
Hamanofim 9, POB 12256
Herzeliya, Israel 46733
Email: ronc@lyciumnetworks.com

Prayson Pate
Overture Networks
P. O. Box 14864

RTP, NC, USA 27709
Email: prayson.pate@overturenetworks.com

Craig White
Level3 Communications, LLC.
1025 Eldorado Blvd,
Broomfield CO 80021
Email: Craig.White@Level3.com

Appendix A. SONET/SDH Rates and Formats

For simplicity, the discussion in this section uses SONET terminology, but it applies equally to SDH as well. SDH-equivalent terminology is shown in the tables.

The basic SONET modular signal is the synchronous transport signal-level 1 (STS-1). A number of STS-1s may be multiplexed into higher-level signals denoted as STS-N, with N synchronous payload envelopes (SPEs). The optical counterpart of the STS-N is the Optical Carrier-level N, or OC-N. Table 2 lists standard SONET line rates discussed in this document.

OC Level	OC-1	OC-3	OC-12	OC-48	OC-192
SDH Term	-	STM-1	STM-4	STM-16	STM-64
Line Rate(Mb/s)	51.840	155.520	622.080	2,488.320	9,953.280

Table 2. Standard SONET Line Rates

Each SONET frame is 125 us and consists of nine rows. An STS-N frame has nine rows and N*90 columns. Of the N*90 columns, the first N*3 columns are transport overhead and the other N*87 columns are SPEs. A number of STS-1s may also be linked together to form a super-rate signal with only one SPE. The optical super-rate signal is denoted as OC-Nc, which has a higher payload capacity than OC-N.

The first 9-byte column of each SPE is the path overhead (POH) and the remaining columns form the payload capacity with fixed stuff (STS-Nc only). The fixed stuff, which is purely overhead, is N/3-1 columns for STS-Nc. Thus, STS-1 and STS-3c do not have any fixed stuff, STS-12c has three columns of fixed stuff, and so on.

The POH of an STS-1 or STS-Nc is always nine bytes in nine rows. The payload capacity of an STS-1 is 86 columns (774 bytes) per frame. The payload capacity of an STS-Nc is (N*87)-(N/3) columns per frame. Thus, the payload capacity of an STS-3c is (3*87 - 1)*9 = 2,340 bytes per frame. As another example, the payload capacity of an STS-192c is 149,760 bytes, which is 64 times the capacity of an STS-3c.

There are 8,000 SONET frames per second. Therefore, the SPE size, (POH plus payload capacity) of an STS-1 is 783*8*8,000 = 50.112 Mb/s. The SPE size of a concatenated STS-3c is 2,349 bytes per frame or 150.336 Mb/s. The payload capacity of an STS-192c is 149,760 bytes per frame, which is equivalent to 9,584.640 Mb/s. Table 2 lists the SPE and payload rates supported.

SONET STS Level	STS-1	STS-3c	STS-12c	STS-48c	STS-192c
SDH VC Level	-	VC-4	VC-4-4c	VC-4-16c	VC-4-64c
Payload Size(Bytes)	774	2,340	9,360	37,440	149,760
Payload Rate(Mb/s)	49.536	149.760	599.040	2,396.160	9,584.640
SPE Size(Bytes)	783	2,349	9,396	37,584	150,336
SPE Rate(Mb/s)	50.112	150.336	601.344	2,405.376	9,621.504

Table 2. Payload Size and Rate

To support circuit emulation, the entire SPE of a SONET STS or SDH VC level is encapsulated into packets, using the encapsulation defined in section 5, for carriage across packet-switched networks.

Full Copyright Statement

Acknowledgement

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