Network Working Group   Alexander ("Sasha") Vainshtein
Israel Sasson
Akiva Sadovski
Internet Draft   Axerra Networks

Expiration Date: Eduard Metz
August 2002   KPNQwest

Tim Frost
Zarlink Semiconductor

February 2002

TDM Circuit Emulation Service over Packet Switched Network (CESoPSN)

draft-vainshtein-cesopsn-02.txt

Status of this Memo

Abstract

This document describes a method for encapsulating TDM digital signals defined in the plesiochronous digital hierarchy (PDH) as a pseudo-wire (PW) over various packet-switched networks (PSN). In this regard this document complements similar work for SONET/SDH circuits.

Proposed PW encapsulation uses RTP for clock recovery and supports signaling between Provider Edge (PE) devices.
Encapsulation proposed in this document may be extended to low-rate SONET/SDH traffic as well.

TABLE OF CONTENTS

1. Introduction

This document describes requirements for edge-to-edge emulation of time division multiplexed (TDM) digital signals defined in Plesiochronous Digital Hierarchy (PDH), see [G.703], [G.704], [T.107] [T1.103] and [T1.107a] and a corresponding encapsulation technique.

To support TDM traffic, which includes voice, data, and private leased line service, the network must emulate the circuit characteristics of a TDM network. A new circuit emulation header and RTP-based mechanisms for carrying clock over PSN are used to encapsulate TDM signals and provide the Circuit Emulation Service

over PSN (CESoPSN).

Primary application of the technique described in this document is emulation of PDH circuits in situations when native PDH traffic is generated by CE devices and does not depend upon the way this traffic reaches PE devices. However, its use may be extended to carrying SDH traffic as "unstructured TDM", thus providing an alternative to the approach defined in [MALIS].

The CESoPSN solution presented in this document fits the framework for PW services as described in [PWE3-FW] and satisfies the general requirements put forward in [PWE3-REQ].

2. Summary of Changes from the -01 Revision

Note: This section will be removed from the final document.

1. A section on generic and service-specific requirements for edge-to-edge emulation of TDM circuits has been added
2. Fractional E1/T1 has been consistently replaced with N*DS0

3. Support of channel-associated CE signaling (CAS) for N*DS0 services based upon the techniques defined in [RFC2833] has been added
4. The structure of the control word has been aligned with the [MARTINI-ENCAP]
5. References have been updated in accordance with the latest developments
6. RTP Payload Types have been decoupled from PW types. Dynamic allocation of PT values will be used instead
7. Most of the text that should logically belong to more generic PWE3 documents and/or tutorials has been removed
8. In-band CESoPSN loopback commands have been removed
9. G.826-compatible PM parameters for CESoPSN have been defined
10. A brief description of adaptive jitter buffer behavior has been added.

3. Terminology and Reference Models

3.1. Terminology

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 [RFC2119].

The terms defined in [PWE3-FW], Section 1.4 are consistently used, usually without additional explanations. However:

- o The terms 'CE-bound' and 'PSN-bound' are consistently used instead of 'outbound' and 'inbound' when describing traffic directions
- o The term "Interworking function" (IWF) is often used for describing the protocol operation with explicit references to CE-bound or PSN-bound direction of the IWF.

Some terms and acronyms are commonly used in conjunction with the TDM services. In particular:

- o Alarm Indication Signal (AIS) is a common term denoting a special bit pattern in the TDM bit stream that indicates presence of an upstream circuit outage
- o Channel-Associated Signaling (CAS) is one of several signaling techniques used by the telephony applications to convey various states of these applications (e.g., off-hook and ob-hook). CAS uses a certain, circuit-specific multiframe structure that is imposed on the TDM bit stream and a predefined association between the relative timeslot (= channel) number within this stream and position of certain bits within this multiframe structure. Up to 16 application states can be distinguished and signaled (see [G.704] for details).

3.2. Reference Models

3.2.1. Generic Models

Generic models that have been defined in Sections 3.1 (Network Reference Model), 3.2 (Maintenance Reference Model), 3.4 (Protocol Stack Reference Model) and 3.5(Logical Protocol Layering Model) of [PWE3-FW] are fully applicable for the purposes of this document without any modifications.

All the services considered in this document represent special cases of the generic circuit-oriented payload type defined in Section 3.5.2.1 of [PWE3-FW].

3.2.2. Synchronization Considerations and Deployment Scenarios

Two basic issues must taken into account regarding possible synchronization techniques for emulation of circuit-oriented services:

- o Can all the PE devices of the given pseudo-wire domain (PWD) be synchronized? Or, in more precise terms, is the same high-quality synchronization source available to all the PE devices in the given PWD?
- o Is the CE device synchronized to the same source as its 'local' PE?

The answer to the first question depends upon design of the specific PSN. E.g. PE devices in a PSN based entirely on POS links can be easily synchronized while PE devices of a PSN based on Gigabit Ethernet links (or on a mix of Gigabit Ethernet and POS) would as often as not remain unsynchronized.

The answer to the second question depends on specifics of the customers served by the PSN operator. In particular, if the CE devices are just nodes in the customers' TDM networks with their own synchronization schemes, they would probably continue to use these schemes even if the PSN is fully synchronized.

Combinations of answers to these basic questions provide at least three viable deployment scenarios:

1. "One Synchronous Network" Scenario, i.e.:
   a. The same high-precision synchronization source is available in all the PE devices of the given PSN
   b. This synchronization source is also used by all the CE devices terminating TDM end services of PWs crossing the PSN
   c. The PW mechanisms must provide compensation only for the packets inter-arrival jitter introduced by the PSN
2. "Synchronous Carriers' Carrier" Scenario, i.e.:
   a. The same high-precision synchronization source is available in all the PE devices of the given PSN
   b. Each Emulated circuit connects two CEs that are either loop-timed to the corresponding PE or synchronized to their own synchronization source

   c. The PW must carry the difference between the PSN clock and the CE clock over the PSN as well as compensate the packets' inter-arrival jitter introduced by the PSN
3. "Asynchronous Carriers' Carrier" Scenario, i.e.:

a. Each PE uses its own synchronization source. The quality
   of this source is selected in accordance with requirements
   of the emulated services (e.g., a Stratum 4 clock is
   sufficient for E1 and T1 services)
b. Each emulated circuit connects two CEs that are either
   loop-timed to the corresponding PE or synchronized to
   their own synchronization source
c. Every direction of the PW must carry the original line
   clock of its end service across the PSN as well as
   compensate for the packets' inter-arrival jitter
   introduced by the PSN.

3.2.3. Service Examples

Fig.1 below presents several examples of a T1 Emulated Service.

```
                           _/_  \    /    \    / \
+-------+ Physical         /+-+            \__/   \     _ Hub Site
|Site A|    T1            / |P| +----+             \      (CE-3)
|T1 #1=|===================|E|=| R |   +---+ +-+   \ OC12+-------+
|(CE-1)|                  \ |1| |    |===|   | | |---------|       |
+-------+                 / +-+ +----+   |   | | | ========|=T1 #1|
                         /               | R |=|P|         |       |
+-------+ T1 +----+  DS3 /   +-+ +----+  |   | |E| ========|=T1 #2|
|Site B|    |    |-----------|P| | R |===|   | |3|---------|       |
|T1 #2=|====|M13|===========|E|=|    |   +---+ +-+    /    +-------+
|(CE-2)|    |    |-----------|2| +----+              /
+-------+   +----+          \ +-+                  /
                             \   ___      ___    /
                              \_/   \____/   \___/
```

Figure 1: T1 Emulation Example Diagram

In this diagram, T1 circuits are attached to the PE devices in three different ways:

- o As a physical T1 line (between CE-1 and PE-1)
- o As a virtual T1 signal multiplexed in DS3 using one of possible multiplexing formats (between CE-2 and PE-2, see [T1.103] for details). M23 is a PDH multiplexor
- o As a virtual T1 signal mapped into an appropriate SONET virtual tributary, the latter being multiplexed in OC-12 (between CE-3 and PE-3 - see [T1.105] or [G.707] for details).

4. Scope and Requirements

4.1. Emulated Services

4.1.1. PDH Circuits

This specification describes service-specific encapsulation layer for edge-to-edge emulation of the following TDM services over a PSN:

1. Structured services:
   a. Transparent N*DS0, 1 <= N <= 31 as described in [G.704].
   b. N*DS0 with channel-associated signaling (CAS) as described in [G.704], 1<= N <= 30
2. Unstructured services
   a. Unstructured E1 as described in [G.704]
   b. Unstructured T1 (DS1) as described in [T.157a]
   c. Unstructured E3 as defined in [G.751]
   d. Unstructured T3 (DS3) as described in [T.157a]

4.1.2. SONET/SDH Circuits

Encapsulation layer described in this specification MAY be, with some modifications, also used for emulation of unstructured "low-rate" (STS-1/STM-0, STS-3c/STM-1) SONET/SDH circuits. Details are discussed in Annex B.

4.2. Scope

This specification defines only the encapsulation layer for edge-to-edge emulation of TDM services mentioned in Section 4.1.

In accordance with the logical protocol layering architecture for PWE3, the encapsulation layer MUST NOT be dependent upon specific instantiations of:

1. The PSN layer (i.e. IPv4, IPv6 or MPLS). In order to satisfy this requirement, encapsulation should be used on packets of fixed size to avoid possible need in the PSN-specific optional length service
2. Multiplexing layer. In order to satisfy this requirement and, at the same time, to allow detection of 'stray packets' the encapsulation header SHOULD provide some means for identifying the packets as belonging to the PW.

4.3. Generic Requirements

Note: This and the following section should be split into a separate requirements document.

4.3.1. Relevant Common PW Requirements

The encapsulation layer for TDM services considered in this document should comply with the following common PW requirements defined in [PWE3-REQ]:

1. Conveyance of Necessary L2/L1 Header Information - relevant only for TDM structured services

2. Support of Multiplexing and Demultiplexing if supported by the native services - relevant for N*DS0 circuits with or without CAS
3. Handling Control Messages of the Native Services - relevant only for structured TDM services

4. Consideration of the PSN Tunnel Header Overhead (see also Section 4.4.4 below)
5. Detection and handling of PW faults (see also Section 4.4.5 below). In particular, ability to detect loss of packets SHOULD be supported in order to allow differentiation between outages of the emulated service resulting from PSN problems and these resulting from problems beyond the PSN
6. Clock Recovery (see also Section 4.4.2 below).

4.3.2. Common Circuit Payload Requirements

All the services considered in this document belong to the generic 'Circuit Payload' type defined in [PWE3-FW], Section 3.5.2.1.1.

Accordingly, the encapsulation layer MUST provide the common Sequencing service and SHOULD provide timing information.

The encapsulation layer for the Circuit Payload services does not necessarily have to provide the length service.

4.3.3. The Principle of Minimal Intervention

The encapsulation layer SHOULD comply with the principle of minimal intervention as described in [PWE3-LAYERS], Section 4.3.5.

4.4. Service-Specific Requirements

4.4.1. Interworking

1. The encapsulation layer MUST support network interworking between end services of the same type and bit-rate.
2. The encapsulation layer SHOULD remain unaffected by specific characteristics of connection between the end services and PE

devices at the two ends of the PW (see service examples in Section 3.2.3 above).

4.4.2. Network Synchronization Schemes

The encapsulation layer MUST be applicable to all the network synchronization schemes mentioned in Section 3.2.2.

If the same high-quality synchronization source is available to all the PE devices in the given domain the encapsulation layer SHOULD be able to infer additional benefits (e.g., facilitate better reconstruction of the native service clock) from this fact.

4.4.3. CE Signaling

Unstructured TDM services do not usually require any special mechanisms for carrying CE signals as these would be carried as part of the emulated service.

Structured TDM services may require application-specific CE signaling.

In some cases this signaling may require synchronization with the data. E.g., code-associated signaling (CAS) reflects the state of telephony applications (like off-hook and on-hook) that must be passed across the emulated service and synchronized with data to allow normal operation of these applications.

The encapsulation layer SHOULD support signaling of state of CE applications for the relevant services providing for:

- o Multiplexing of application-specific CE signals and data of the emulated service in the same PW
- o Synchronization (within the application-specific tolerance limits) between CE signals and data at the PW egress
- o Probabilistic recovery against possible accidental loss of signaling packets in the PSN
- o Deterministic recovery of the CE application state after PW setup and network outages.

Some types of CE signaling associated with the TDM circuits (e.g., performance monitoring requests and responses, requests to operate and release loopbacks etc.) do not reflect application state and

hence do not require synchronization with data. As a consequence, these signals can be passed out-of-band and do not have to be supported by the encapsulation layer.

The payload format for the 'signaling' packets MAY be application-specific.

4.4.4. Latency and Encapsulation Effectiveness

The encapsulation layer SHOULD allow for an effective trade-off between the following requirements:

1. Effective PSN bandwidth utilization. Assuming that the size of encapsulation layer header does not depend on the size of its payload, increase in the packet payload size results in increased efficiency.
2. Low edge-to-edge latency. Low end-to-end latency is the common requirement for Voice applications over TDM services. Packetization latency is one of the components comprising edge-to-edge latency and decreases with the packet payload size.

4.4.5. Fault Detection and Handling

The encapsulation layer for edge-to-edge emulation of TDM services SHOULD, separately or in conjunction with the lower layers of the pWE3 stack, provide for detection of the following defects:

1. Misconnection
2. Loss of packets. Special importance of detection of this defect has been explained in Section 4.3.1 above
3. Malformed packets
4. Loss of synchronization.

4.4.6. Performance Monitoring

The encapsulation layer for edge-to-edge emulation of TDM services should provide for collection of performance monitoring (PM) data that is compatible with the parameters defined for 'classic', TDM-based carriers of these services (see [G.826] for details).

4.4.7. Bandwidth Saving

The encapsulation layer should provide for saving the PSN bandwidth

by not sending invalid data.

4.4.8. Adaptation of the Jitter Buffer

The encapsulation layer SHOULD allow adaptation of the jitter buffer size to the actually observed level of the packets' inter-arrival jitter while maintaining acceptable levels of errors that are introduced by such an adaptation.

Note: The meaning of 'acceptable level of errors' depends on the application using the emulated service. In particular, Voice applications can tolerate loss or insertion of a single octet in a contiguous sequence of several non-erroneous octets. (In case of insertion, it is customary to repeat the previous, non-erroneous, octet.)

5. CESoPSN Encapsulation

5.1. Generic CESoPSN Format

CESoPSN packets use format shown in Fig. 2 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
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |                              ...                              |
   |              PSN and multiplexing layer headers               |
   |                              ...                              |
   +=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+
   |                           Fixed                               |
   +--                                                           --+
   |                            RTP                                |
   +--                                                           --+
   |                     Header (see [RFC1889])                    |
```

```
+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+
|                CESoPSN Control Word (optional)                |
+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+
|            Packetized TDM data or CE signaling data           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 2. CESoPSN Format

5.2. CESoPSN Header

The CESoPSN header includes a fixed RTP header (12 octets) and an optional CESoPSN Control Word (4 octets).

5.2.1. Usage of RTP Header

CESoPSN uses the fields of the fixed RTP header (see [RFC1889], Section 5.1) in the following way:

- 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 to for CESoPSN packets carrying PDH circuits. CESoPSN packets carrying unstructured SONET/SDH circuits MAY set this bit to 1 to distinguish packets that carry the framing octets
- o PT (payload type) is used to distinguish between packets carrying the packetized TDM data and packets carrying CE signaling. At least one PT value should be allocated from the range of dynamic values (see [RTP-TYPES]) for every CESoPSN 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. The PE at the PW egress MAY use this value to detect malformed packets. An additional PT value from the same range MUST be allocated for CESoPSN PWs supporting in-band CE signaling (see Section 5.3.2 below)
- 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 a multiple of 8 KHz. Possible modes of timestamp generation are discussed below

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

Note: The same PT value can be safely allocated for different PWs.

The RTP header in CESoPSN can be used in conjunction with at least the following modes of timestamp generation:

1. Absolute mode: the ingress PE sets time stamps using the clock recovered from the incoming TDM bit stream
2. Differential mode: PE devices connected by the PW have access to the same high-quality synchronization source, and this synchronization source is used for timestamp generation.

Usage of other timestamp generation modes is left for further study.

Absolute mode allows operation in the Asynchronous Carrier's Carrier deployment scenario. Differential mode may improve quality of the recovered clock in the One Synchronous Network and Synchronous Carrier's Carrier deployment scenarios.

5.2.2. Usage and Structure of the Control Word

Usage of the CESoPSN control word allows:

- o Differentiation between the PSN problems and the problems beyond the PSN as causes for the emulated service outages
- o Saving bandwidth by not transferring invalid data (AIS, idle code)
- o Signaling problems detected at the PW egress to its ingress

Consequently, usage of the CESoPSN Control Word is the recommended default. The PE peers MAY agree not to use it in a specific CESoPSN PW as part of the PW setup process.

Note: Alternative techniques for conveying forward and backward indications without using the control word are left for further study.

The structure of the CESoPSN Control Word is shown in Fig. 3 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|0|0|0|0|A|I|L|T|Z|            Reserved                         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 3. Structure of the CESoPSN Control Word

- o Bits 0-3 MUST be set to 0 at ingress and MUST be ignored at egress
- o Bit A - carries Local AIS indication. If set, represents AIS of the carried unstructured circuit. A packet with the A bit set MAY carry no payload
- o Bit I - carries Local Idle Code indication. If set, represents the Idle Code in the payload of a N*DS0, a N*DS0 with CAS or an unstructured T3 circuit. A packet with the I bit set MAY carry no payload
- o Bit L - carries Remote Loss of Packets indication of the PW carrying CESoPSN, i.e., this bit is set in packets transmitted by PE-2 to PE-1 if PE-2 detected loss of packets in the stream received from PE-1
- o Bit T - carries Remote Synchronization Problem indication.
- o Bit Z - if set, indicates that the CESoPSN IWF operates under a PW loopback command (regardless of the origin of this command). If cleared, indicates normal CESoPSN IWF operation
- o Reserved - these bits are reserved for possible future use. Currently they MUST be set to 0 at ingress and ignored at egress.

Notes:

1. Either A or I bit (but not both) can be set in the CESoPSN control word.
2. Information about lost packets (carried via the L bit) can be used at ingress as an indication to resynchronize CE application state, see Section 5.3.2 below.

5.3. Payload Data Format

A single CESoPSN packet always contains one or more native circuit frames of the carried circuit. This provides for emulation of performance monitoring parameters of "classic" carriers of TDM circuits (e.g., SONET/SDH).

Note: The native circuit frames for all the circuits considered in this document save from unstructured T1 are octet-aligned. The T1 native circuit frame (193 bits) is not, and hence requires special treatment - see Section 5.3.4 below.

The PSN operator selects the number of native service frames in a CESoPSN packet for a specific PW taking into account the following considerations:

- o Packetization latency requirements vs. bandwidth utilization (see Section 4.4.4 above)

o Path MTU limitations in order to avoid fragmentation of CESoPSN packets

This specification assumes that the number of native service frames in a CESoPSN packet is:

o Defined during the PW setup and remains constant for the duration of a PW. Such an arrangement simplifies implementation because it implies that the CESoPSN packets are transmitted at a constant rate

o The same for both directions of the PW. Such an arrangement simplifies signaling and processing of backwards problem indications.

5.3.1. Transparent N*DS0 Circuits

The payload data format for transparent N*DS0 circuits is shown in Fig. 4 below (N - number of timeslots in the circuit, M = number of the native circuit frames in a CESoPSN packet, the 1st timeslot of the 1st native frame is the 1st octet of the payload). The matrix shown in this diagram is mapped into array of payload octets row by row.

```
Timeslots ->|      1       |     2        | ... |       N      |
------------+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+
 N  C  F   1|              |              | ... |              |
            +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 a  i  r   2|              |              | ... |              |
            +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 t  r  a ...|              |              | ... |              |
            +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 i  c  m ...|              |              | ... |              |
            +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 v  u  e ...|              |              | ... |              |
            +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 e  i  s ...|              |              | ... |              |
            +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    t      M|              |              | ... |              |
            +=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+
```

Figure 4. Payload structure for a N*DS0 Circuit

CESoPSN-based emulation of a transparent N*DS0 TDM circuit can be considered as "bundling" of N independent DS0 circuits (see [PWE3-REQ], Section 2.1.3).

The payload structure described provides for adaptation of the jitter buffer size for Voice applications while maintaining acceptable level of errors:

- o Actual size of the jitter buffer can be decreased by "shortening" the payload of some of the packets already in the buffer by the one "row" (native circuit frame) when they are transmitted. This is equivalent to dropping one octet from each timeslot
- o Actual size of the jitter buffer can be increased by "lengthening" the payload of some of the packets already in the buffer by one "row" (native circuit frames) when they are transmitted. This is equivalent to insertion of a single octet into each timeslot; the values carried in the last actual row of the matrix are repeated.

5.3.2. N*DS0 circuits with CAS

A PW that emulates an N*DS0 circuit with CAS assumes that CE devices are PSTN switches that synchronize the state of each of N DS0 channels using channel-associated signaling. This PW carries TDM data in format described in the previous section.

In addition, it carries the CAS state vector of each CE in special signaling packets using:

- o An additional PT value allocated for this purpose from the range of unused values (see [IANA]). This value MUST be different from one allocated for the TDM data packets for the same PW
- o An additional SSRC value that MUST be different from one used for the data packets in order to allow a separate numbering sequence for the signaling packets
- o A sequence numbering scheme that does not depend on one used for the data packets. This allows re-use of common sequence numbers-based mechanisms (like reordering and detection of lost packets) for the data packets for all types of circuits
- o The signaling payload format described in Fig. 5 below. Format of the 32-bit timeslot signaling word is defined in [RFC2833] Section 3.5 and Section 3.14, and numbering of timeslots corresponds to that of the "columns" in the data packets' payload, see Fig. 4.

```
      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
     +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
     |              Timeslot signaling word for TS-1                 |
     +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
     |              Timeslot signaling word for TS-2                 |
     +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
     |                           ...                                 |
     +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
     |              Timeslot signaling word for TS-N                 |
     +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 5. Payload of a Signaling Packet for a N*DS0 Circuit with CAS

Note: The "volume" field defined in the [RFC2833] Section 3.5 is not used with CAS events.

CESoPSN does not require handling of loss of signaling packets; as a consequence, detection of loss of these packets is not required either. On the other hand, the same synchronization source MUST be used for timestamps in both signaling and data packets in order to synchronize data and signaling within reasonable limits.

Signaling packets are generated by the ingress PE in accordance with the following logic (adapted from [RFC2833]):

1. The CESoPSN signaling packet with the same information is sent 3 times at an interval of 5 ms under one of the following conditions:
   a. The CESoPSN PW has been set up
   b. A change in CAS state of one of the timeslots has been detected. If another change of CAS state has been detected during the 15 ms period, this process continues
   c. Loss of packets defect has been cleared
   d. Remote Loss of Packets indication has been cleared (after previously being set)
2. Otherwise, the CESoPSN signaling packet with the current CAS state information is sent every 5 seconds.

These rules allow fast probabilistic recovery after loss of a single

signaling packet as well as deterministic (but, possibly, slow) recovery following PW setup and PSN outages.

5.3.3. Unstructured TDM Circuits

Basically, unstructured TDM circuits do not require framers in the PE devices, and are transferred as bit streams. However, presence of a framer allows detection of some outages of the end services. As a consequence, efficiency of the CESoPSN operation under such outages may be increased.

The payload of a CESoPSN packet carrying an unstructured TDM circuit with an octet-aligned native circuit frame MUST contain one or more native circuit frames of the carried circuit, but no alignment with the framing structure of the service is required.

5.3.3.1 "T1-in-E1" Mode for Unstructured T1 Circuits

As mentioned above, unstructured T1 represents the only case of a TDM circuit considered in this document with a non-octet aligned native circuit frame. In order to accommodate this type of circuit into the general CESoPSN framework, a special "T1 in E1" payload format (similar to one defined in [G.802]) is used as shown in Fig 5 below (M = number of native frames in the CESoPSN packet, D denotes the payload data bits).


```
"Timeslots" |      1         | ... |     24         |      25        |
            |0 1 2 3 4 5 6 7| ... |0 1 2 3 4 5 6 7|0 1 2 3 4 5 6 7|
------------+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+
 N  C  F   1|D D D D D D D D| ... |D D D D D D D D|D|  padding     |
            +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 a  i  r   2|D D D D D D D D| ... |D D D D D D D D|D|  padding     |
            +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 t  r  a ...|D D D D D D D D| ... |D D D D D D D D|D|  padding     |
```

```
           +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 i  c  m ...|D D D D D D D D| ... |D D D D D D D D|D|  padding    |
           +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 v  u  e ...|D D D D D D D D| ... |D D D D D D D D|D|  padding    |
           +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 e  i  s ...|D D D D D D D D| ... |D D D D D D D D|D|  padding    |
           +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    t      M|D D D D D D D D| ... |D D D D D D D D|D|  padding    |
           +=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+
```

Figure 6. The "T1-in-E1" CESoPSN Payload Format

Note: Each row in the matrix presented in Fig. 6 contains exactly 193 payload data bits (and 7 padding bits). However, no alignment of the rows with the T1 framing structure is implied and hence support of this mode does not require a T1 framer in PE.

6. CESoPSN Operation

Note: This section includes non-normative information and implementation considerations. These elements will be moved to an appropriate Appendix in the next update.

Edge-to-edge circuit emulation of a TDM circuit using CESoPSN assumes the following elements:

- o Two PW end services of the same type and bit rate
- o Packetizer at the PW ingress
- o Jitter buffer and de-packetizer at the PW egress.

Setup of a CESoPSN PW assumes exchange of the following information:

- o Types of end services. In order to be connected by a CESoPSN PW, these types MUST be the same and define the PW type. PW types supported by CESoPSN MUST be accommodated into the common enumeration of PW types
- o Bit rates of end services. In order to be connected, bit rates of the two end services MUST be the same and define the PW bit rate
- o Encapsulation layer-specific parameters that define specific instantiation of the protocol

This document defines how the values of these parameters should be encoded. The actual signaling protocols for exchanging these

parameters between the PE peers ("PE/PW signaling" in terms of [PWE3-FW]) are out of scope of this document.

Description of the CESoPSN-based edge-to-edge circuit emulation includes the following elements:

- o Definition of the end service inactive state behavior towards the CE
- o Description of the IWF operation in CE-bound and PSN-bound direction.

Details are presented below.

6.1. Payload Parameters

6.1.1. PW Type

PW types (a.k.a. VC types) have been defined in [MARTINI-TRANS]. PW types used for CESoPSN PW are assigned in such a way as to avoid overlap with types assigned in other PWE3 documents.

The following PW types are defined in this document for CESoPSN-based PWs:

- o Transparent N*DS0 - 65
- o N*DS0 with CAS - 66
- o Unstructured E1 - 67
- o Unstructured T1, bit stream mode - 68 (not defined in this specification)
- o Unstructured T1, T1-in-E1 mode - 69
- o Unstructured E3 - 70
- o Unstructured T3 - 71
- o Unstructured SONET/SDH - 72 (see Annex B).

6.1.2. Circuit Bit Rate

The circuit bit rate is encoded as the number of "timeslots" in the matrix structure of the corresponding CESoPSN data packet.

The following values are used:

- o Transparent N*DS0 - N, 1 <= N <= 31
- o N*DS0 with CAS - N, 1 <= N <= 30
- o Unstructured E1 - 32
- o Unstructured T1, T1-in-E1 mode - 25
- o Unstructured E3 - 537
- o Unstructured T3 - 699
- o Unstructured STS-1 - 810
- o Unstructured STM-1 - 2430

Note: N*DS0, unstructured E1 and unstructured T1 circuits can be carried over any PSN implementing the minimal MTU as defined in [RFC1122]. Unstructured E3 and T3 can be carried over any PSN providing Path MTU of 1.5 Kbytes. Unstructured STS-1 and STM1 are

considered in Annex A.

6.2. Encapsulation Layer Parameters

6.2.1. Usage of Control Word

TRUE value (default) of this Boolean parameter means that the CESoPSN control word is used.

CESoPSN MAY allow negotiation of this parameter, so that the control word will not be used if both sides agree to that.

6.2.2. RTP Payload Type

1. One PT value MUST be allocated from the range of dynamically allocated payload types for each CESoPSN PW for use in the data packets:
   a. The same value MUST be allocated for both directions of the PW
   b. Ingress PW MUST set the PT in the RTP header of all the data packets to the allocated value
   c. Egress PW MAY use this value to detect non-data PW packets. These packets can be either relegated to signaling or considered as malformed
2. For emulation of a N*DS0 circuit with CAS, an additional PT value MUST be allocated from the range of dynamically allocated payload types for each CESoPSN PW for use in the data packets:
   a. It MUST be different from the PT value allocated for data packets
   b. The same value MUST be allocated for both directions of the PW
   c. Ingress PW MUST set the PT in the RTP header of all the signaling packets to the allocated value
3. Egress PW MAY use this value to distinguish signaling PW packets.

Note: The same PT value may be allocated for multiple PWs.

6.2.3. Payload Bytes

This parameter has been defined in [MARTINI-TRANS]. In order to establish a CESoPSN-based PW, the following conditions MUST be met:

o The number of payload bytes MUST be the same for both directions of the PW
o The number of payload bytes MUST be a multiple of the encoded Circuit Bit Rate (see Section 6.1.2 above). E.g., the value of this parameter for an Unstructured E1 circuit (Circuit Bit Rate = 32) with M native circuit frames packet into a single CESoPSN packet will be 32*M, while for an Unstructured T1 it will be 25*M

o The size of the resulting PW packet (including all the headers) SHOULD NOT exceed the path MTU between the participating PEs as provided by the Carrier layer.

Note: For N*DS0 with CAS circuits this parameter defines the number of payload bytes in the data packets only. The number of payload bytes in the signaling packets is inferred from the encoded circuit bit rate in the obvious way.

6.2.4. Timestamp Resolution

This parameter encodes the rate of the clock used for setting timestamps in RTP headers as a multiple of the basic 8 KHz rate.

6.2.5. Synchronization Source ID

The same 32-bit SSRC value MUST be assigned to all the data packets of a given direction of a CESoPSN PW. The CE-bound direction of the IWF MAY be use this value for misconnection detection, especially if such a service is not provided by the PSN and/or multiplexing layer(s).

If data and signaling packets are multiplexed in the same PW, the signaling packets MUST use a separate SSRC value. This arrangement complies with the RTP specification [RFC 1889] and allows effective compression of the PW headers by the standard compressors.

6.2.6. Timestamp Generation Mode

This parameter accepts at least the following two values corresponding to operation modes described in Section 5.2.1:

- o Absolute (1)
- o Differential (2).

6.3. End Service Inactivity Behavior

While the PW is inactive:

- o Each unstructured end service MUST send AIS to its prospective CE
- o Each structured end service MUST send an appropriate Idle Code to its prospective CE

6.4. Description of the IWF operation

Once the PW is set up, the CESoPSN IWF operates like following:

6.4.1. PSN-bound Direction

1. End service data is packetized in accordance with the number of payload bytes specified. For N*DS0 services, the packetized data are aligned with the native circuit frames as described in Section 5.3.1
2. Sequence numbers and timestamps representing the selected synchronization clock are inserted in the CESoPSN headers

3. CESoPSN, multiplexing and PSN headers are prepended to the packetized circuit data
4. Resulting packets are transmitted via the PSN
5. If the PE detects any outage of the incoming an unstructured end service that natively would result in sending the "downstream AIS", the CESoPSN IWF using the control word MUST set the local AIS indication flag (bit A) in the control word. The packet payload MAY be omitted in order to save the PSN bandwidth.
6. If the PE detects an Idle Code condition of the incoming an unstructured T3 end service, or an AIS-producing condition is detected in the incoming 'carrier service' of an N*DS0 end service, the CESoPSN IWF using the control word MUST set the local Idle Code indication flag (bit I) in the control word. The packet payload MAY be omitted in order to save the PSN bandwidth.

Local AIS and Idle Code indications in the CESoPSN control word provide for the following functionality:

- o Ability to distinguish between the PSN problems and ones beyond the PSN as causes of outages of the emulated service
- o Ability to save the PSN bandwidth (but not its switching capacity) by not sending invalid data across the PSN.

The techniques to save the PSN switching capacity in case of an end service outage are left for further study.

6.4.2. CE-bound Direction - Normal Operation

1. The CE-bound IWF includes a jitter buffer that accumulates data from incoming CESoPSN packets with their respective timestamps. The length of this buffer SHOULD be configurable to allow adaptation to various network delay behavior patterns. Size of the jitter buffer is a local parameter of the CESoPSN IWF. Since any CESoPSN data packet carries a fixed number of native data frames of the emulated service, the jitter buffer can be considered as a matrix with "rows" corresponding to native service frames, too.
2. Initially the Jitter buffer is filled with the appropriate inactivity (AIS or Idle) code.
3. Immediately after start, IWF:
   a. Begins reception of incoming CESoPSN packets. PSN and multiplexing layer headers are stripped from the received packets, and packetized TDM data from the received packets is stored in the jitter buffer
   b. Continues to play out its appropriate inactivity code into its end service as long as the jitter buffer has not yet accumulated sufficient amount of data
   c. Signals the CE-bound direction of the local IWF to transmit CESoPSN packets with the T bit set (if control word is used)
4. Once the jitter buffer contains sufficient amount of data (usually half of its capacity), the IWF starts replay of this

   data in its end service in accordance with its (locally defined) 8 KHz transmission clock, so that a single "row" of the jitter buffer matrix is replayed per "tick" of the clock. At the same moment it signals the PSN-bound direction of IWF to clear the T bit in the CESoPSN packets it transmits (if the control word is used)
5. If transmission clock must be recovered from the PW, the timestamps of data packets SHOULD be used for correcting initial transmission clock frequency in accordance with the specified mode of their generation.
6. If adaptation of the jitter buffer size is implemented, it SHOULD NOT introduce additional wander of the transmission clock. It MAY introduce additional errors (e.g., in accordance with the techniques described in Section 5.3.1 above)

7. The CE-bound direction of the IWF:
    a. Performs detection, correlation and handling of CESoPSN faults as described in Section 6.5 below
    b. Collects the PW Performance Monitoring data as defined in Section 6.6 below
8. CE application state signals received in the signaling packets SHOULD be synchronized with data using the timestamps and inserted (in an appropriate format) into the CE-bound TDM stream. Signals that cannot be inserted into the CE-bound TDM stream due to the local format limitations MUST BE ignored. Any aspects of translation of values of CE signals are out of scope of this specification.

6.4.3. IWF Loopback

An IWF loopback for the CESoPSN IWF MAY be set and cleared by an external (management) command.

Once such a loopback is set, the IWF will loop packets coming from the PSN back to the PSN. In addition it will mark these packets by setting Z bit in the CESoPSN control word.

Once the loopback is cleared, the IWF resumes its normal operation.

6.5. CESoPSN Defects

6.5.1. Misconnection

Some combinations of PSN and multiplexing layers (see Annex A) inherently provide for detection of packets that do not belong to the PW ('stray packets').

CESoPSN MAY use the SSRC field in the RTP header for detection of 'stray packets' even if such a capability is provided by the specific combination of PSN and multiplexing layers.

Regardless of the way in which a stray packet has been detected:
- o It MUST be discarded by the CE-bound IWF
- o A counter of 'stray packets' must be incremented

- o If reception of stray packets persists, the Misconnection alarm should be reported to the management system.

The IWF mechanisms for detection of lost packets (e.g., expected next sequence number) MUST NOT be affected by reception of 'stray packets'.

6.5.2. Re-Ordering and Loss of Packets

CESoPSN implementations SHOULD use sequence numbers in the RTP header and expected rate of transmission of data packets for detection of our-of-order delivery and packets' loss. In particular, they MAY maintain the next expected sequence number value that would be:

- o Advanced every time a packet belonging to this PW with an equal or greater (mod 65536) sequence number has been received or a timeout defined by the expected packet arrival rate has expired
- o Used as the center of a sliding window for packet reordering. The size of this window SHOULD be limited by the size of the jitter buffer.

Out-of-order packets that cannot be reordered MUST be considered as lost.

If loss of one or more CESoPSN packets has been detected at the egress of the CESoPSN PW, its jitter buffer MUST be filled with the appropriate amount of the AIS (or Idle - depending on the service type) code to be replayed into the relevant PWES. In addition:

- o If the CESoPSN control word is used, the Remote Lost Packets Indication flag (bit L) MUST be set in the next packet to be sent in the opposite direction of the PW
- o A counter of lost packets must be incremented
- o If the loss-of-packets condition persists, an alarm should be sent to the management system.

6.5.3. Malformed Packets

CESoPSN PW detects a malformed packet using the following rules:

- o The PT value in its RTP header does not correspond to one of the PT values allocated for this PW
- o The actual packet payload size can be unambiguously inferred from the data link, PSN or multiplexing layer of the PW and does not match the payload size defined for the packets of this type in this PW.

If a malformed in-order packet has been received at the egress of a CESoPSN PW, then:

- o Its jitter buffer MUST be filled with the appropriate amount of the AIS (or Idle) code replay to be replayed into the relevant PWES
- o A counter of malformed packets must be incremented
- o If the payload mistype condition persists, an appropriate alarm should be sent to the management system.

6.5.4. Loss of Synchronization

The CESoPSN IWF MAY detect two types of loss of synchronization errors:

6.4.5.1 Jitter Buffer Overrun

This fault is detected if the jitter buffer at the PW egress cannot accommodate the newly arrived CESoPSN packet in its entirety.

A CESoPSN packet that cannot be stored in the jitter buffer MUST be discarded.

If the jitter buffer overrun condition persists, an appropriate alarm should be sent to the management system. In addition, the Remote Loss of Synchronization (bit T) flag SHOULD be set in the next packet to be send in the opposite direction of the service.

6.5.4.2. Jitter Buffer Underrun

This fault is detected if the jitter buffer at the PW egress becomes empty before arrival of a new CESoPSN packet while loss of packets has not been detected. CESoPSN implementations MAY never detect the Jitter Buffer Underrun condition if their packets' loss detection mechanisms do not allow it.

If the jitter buffer underrun condition persists, an appropriate alarm should be sent to the management system. In addition, the Remote Loss of Synchronization (bit T) flag SHOULD be set in the next packet to be send in the opposite direction of the service.

6.6. Performance Monitoring

6.6.1. Errored Data Blocks

[G.826] defines the concept of an errored data block that serves as the basis of for collection of performance monitoring parameters. It also defines the size of the data block for most TDM circuits. These definitions are aligned with the 'native circuit frame' size of these circuits so that every G.826-compatible data block contains an integer multiple of native circuit frames, e.g.:

- o For E1 and T1 circuits, a data block contains 4 native service frames
- o For E3 and T3 circuits, a data block contains one native service frame etc.

The following definitions of error events and errored data blocks for CESoPSN provide for collection of [G.826]-compatible performance

monitoring parameters:

- o An error event is insertion of a single native service frame of inactivity code into the jitter buffer if it does not stem from receiving a CESoPSN packet with an AIS or Idle Code indication

- o An errored data block is a data block defined in accordance with [G.826] that has experienced at least one error event.

6.6.2. Errored, Severely Errored and Unavailable Seconds

The definition of an errored data block presented above can be used to define Errored Seconds, Severely Errored Seconds and Unavailable Seconds in accordance with [G.826].

6.7. QoS Issues

If the PSN providing connectivity between PE devices is Diffserv-enabled and implements EF PHB (see [RFC2598bis]), all the CESoPSN data packets should be marked for EF PHB at ingress. Such an arrangement results in decrease of the packets' inter-arrival jitter and hence in decrease of latency introduced by the TDM circuit emulation.

7. RTP Payload Format Considerations

In accordance with guidelines specified in [RFC2736], the following issues are addressed by this specification:

7.1. Resilience to moderate loss of individual packets

The impact of loss of an individual data packet may be decreased by decreasing the packet size (with the associated loss of efficiency).

Resilience to loss of an individual signaling packet is provided for by the rules described in Section 5.3.2 above.

7.2. Ability to interpret every single packet

This requirement is met since every CESoPSN packet carries a multiple of the native frame of the carried service.

7.3. Non-usage of the RTP Header Extensions

This recommendation is met, since RTP-wise, the CESoPSN Control Word

is part of the RTP payload. Alignment with this requirement facilitates usage of standard header compression mechanisms if CESoPSN uses UDP/IP as its PSN and multiplexing layers.

7.4. Compression of RTP headers

Existing relevant standards ([RFC2508], [RFC3095]) deal with compression of RTP/UDP/IP headers on specific P2P links. Compression techniques defined in these documents are fully applicable for CESoPSN if it uses UDP/IP as PSN and multiplexing layers respectively. Standard compression of CESoPSN/UDP/IP headers will be very effective, since:

- o Value of the SSRC field in the CESoPSN header of data packets remains constant for the duration of a CESoPSN session

- o Value of the Timestamp field in the CESoPSN header is usually incremented by a fixed value from packet to packet
- o CESoPSN control word is NOT defined as RTP header extension.

As a consequence, a PSN-independent end-to-end compression technique of RTP headers seems not justified.

8. Congestion Control (RFC 2914) Conformance

CESoPSN PWs carry constant bit rate (CBR) services. These services, by definition, cannot behave in a TCP-friendly manner prescribed by [RFC2914] under congestion while retaining any value for the user.

Devices implementing CESoPSN and using IP as their PSN layer:

- o MUST set the ECN bits of the IP header (see [RFC3168]) to non-ECT ('00') value at ingress (to prevent routers in the network from setting them to the CE ('11') value
- o SHOULD ignore these bits at egress.

9. FFS Issues

Note: This section will be removed from the final revision of the document.

The following issues will be addressed in the next revisions of this document:

- o Techniques for saving the PSN switching capacity when the PW experiences an end service outage or does not carry any valid data

o Usage of RTCP. One particular application to be considered is retrieval of remote problems' indications without the control word
o Effect of timestamp resolution on quality of clock recovery in Differential mode.

10. Security Considerations

This document does not affect the underlying security issues of specific PSN.
In addition, it defines misconnection detection capabilities of CESoPSN. These capabilities increase resilience of CESoPSN to misconfiguration and some types of DoS attacks.

11. Applicability Statement

CESoPSN is an encapsulation layer intended for carrying TDM circuits (transparent N*DS0, transparent N*DS0 with CAS, unstructured E1/T1 and unstructured E3/T3) over PSN.

Applicability of CESoPSN MAY be extended to low-rate SONET/SDH circuits with minimal modifications.

CESoPSN allows carrying both data and clock of TDM circuits across multiple types of PSN.

CESoPSN allows carrying CE signaling that requires synchronization with data (e.g., channel-associated signaling (CAS) for Voice applications) in-band in separate signaling packets. The RTP Payload Type (PT) is used to distinguish between data and signaling packets, while the Timestamp field is used for synchronization. This makes CESoPSN extendable to support different types of CE signaling without affecting the data path in the PE devices.

CESoPSN does not presume availability of a global synchronous clock at the ends of a PW. This makes it suitable for Asynchronous Carriers' Carrier applications.

CESoPSN uses RTP for carrying the clock across the PSN. The additional CESoPSN header (if used) is a payload format header and hence standard header compression techniques for RTP/UDP/IP profile over links slow and/or error-prone links are fully applicable to CESoPSN PWs.

CESoPSN allows the PSN bandwidth conservation by carrying only AIS and/or Idle Code indications instead of data.

Being a constant bit rate (CBR) service, CESoPSN cannot provide TCP-friendly behavior under network congestion.

CESoPSN allows collection of TDM-like faults and performance monitoring parameters hence emulating 'classic' carrier services of TDM circuits (e.g., SONET/SDH). Similarity with these services is increased by the CESoPSN ability to carry 'far end error' indications.

CESoPSN provides for a carrier-independent ability to detect misconnections and malformed packets. This feature increases resilience of the emulated service to misconfiguration and DoS attacks.

CESoPSN provides for detection of lost packets and hence allows to distinguish between the PSN problems and ones beyond the PSN as causes of outages of the emulated service.

Faithfulness of a CESoPSN PW may be increased if the carrying PSN is Diffserv-enabled and implements EF PHB.

CESoPSN does not provide any mechanisms for protection against PSN outages. As a consequence, resilience of the emulated service to such outages is defined by the PSN behavior. On the other hand, the jitter buffer and packets' reordering mechanisms associated with CESoPSN increase resilience of the emulated service to fast PSN rerouting events.

12. IANA Considerations

This specification requires assignment of new PW Types for CESoPSN PWs as described in Section 6.1.

13. Intellectual Property Considerations

This document is being submitted for use in IETF standards discussions. Axerra Networks, Inc. has filed one or more patent applications relating to the CESoPSN technology outlined in this document. Where there is a necessary dependence upon such patents and patent applications in implementing an IETF adopted standard

resulting from this document, Axerra Networks will license on fair, reasonable, and non-discriminatory terms to all parties, any patent claims it owns covering such technology, solely to the extent such technology is essential to comply with such standard. Any such license to a party shall start on the date that Axerra Networks and the party enter into an agreement related thereto and shall be granted on the condition that any such party grants to Axerra Networks and its corporate affiliates a reciprocal license under such party's patents for which there is also a necessary dependence.

ACKNOWLEDGEMENTS

We express deep gratitude to Stephen Casner who reviewed this document in detail, corrected some serious errors and provided many valuable inputs. Some of his inputs will be explored in the next revisions of the draft.

We thank Sim Narasimha and Yaron Raz for valuable feedbacks.

We thank Alik Shimelmits for many fruitful discussions.

REFERENCES

[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

[PWE3-LAYERS], Stewart Bryant et al., Protocol Layering in PWE3, Work in Progress, February 2002, pwe3-protocol-layering-01.txt
[MALIS] Andrew G. Malis et al, SONET/SDH Circuit Emulation Service Over MPLS (CEM) Encapsulation, Work in progress, April 2001, draft-malis-sonet-ces-mpls-04.txt

[PWE3-SONET] Andrew G. Malis et al, SONET/SDH Circuit Emulation over Packet (CEP), Work in progress, September 2001, draft-malis-pwe3-sonet-00.txt

[KOMPELLA] MPLS-based Layer 2 VPNs, Work in Progress, July 2001, draft-kompella-ppvpn-l2vpn-00.txt

[MARTINI-TRANS] Luca Martini et al, Transport of Layer 2 Frames Over MPLS, Work in progress, November 2001, draft-martini-l2circuit-trans-mpls-08.txt

[MARTINI-ENCAP] Luca Martini et al, Encapsulation Methods for Transport of Layer 2 Frames Over MPLS, Work in progress, November 2001, draft-martini-l2circuit-encap-mpls-04.txt

[L2TPv3] J.Lau et al, Layer Two Tunneling Protocol "L2TP", Work in progress, October 2001, draft-ietf-l2tpext-l2tp-base-01.txt

[RFC1122] R. Braden (ed.), Requirements for Internet Hosts -- Communication Layers, RFC 1122, IETF, 1989

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

[RFC2119] S.Bradner, Key Words in RFCs to Indicate Requirement Levels, RFC 2119, IETF, 1997

[RFC2434] T. Narten, H. Alvestrand, Guidelines for Writing an IANA Considerations Section in RFCs, RFC 2434, IETF, 1998

[RFC2474] K. Nichols et al., Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers, RFC 2474, IETF, 1998

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

[RFC2736] M. Handley, C. Perkins, Guidelines for Writers of RTP Payload Format Specifications, RFC 2736, IETF, 1999

[RFC2598bis] Bruce Davie (ed.), An Expedited Forwarding PHB, Work in Progress, April 2001, draft-ietf-diffserv-rfc2598bis-01.txt

[RFC2833] H. Schulzrinne, S. Petrack, RTP Payload for DTMF Digits, Telephony Tones and Telephony Signals. RFC 2833, IETF, 2000

[RFC2914] S. Floyd, Congestion Control Principles, RFC 2914, IETF, 2000

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

[RFC3140] D. Black et al, Per Hop Behavior Identification Codes, RFC 3140, IETF, June 2001

[RFC3168] K. Ramakrishnan, S. Floyd, D. Black, The Addition of Explicit Congestion Notification (ECN) to IP, RFC 3168, IETF, 2001

[RTP-TYPES] RTP PARAMETERS, http://www.iana.org/assignments/rtp-parameters

[G.704] ITU-T Recommendation G.704 (10/98) - Synchronous frame structures used at 1544, 6312, 2048, 8448 and 44 736 Kbit/s hierarchical levels

[G.707] ITU-T Recommendation G.707 (10/00) - Network Node Interface for Synchronous Digital Hierarchy (SDH)

[G.751] ITU-T Recommendation G.751 (11/88) - Digital multiplex equipments operating at the third order bit rate of 34 368 Kbit/s and the fourth order bit rate of 139 264 Kbit/s and using positive justification

[G.802] ITU-T Recommendation G.802 (11/88) - Interworking between networks based on different digital hierarchies and speech encoding laws

[G.826] ITU-T Recommendation G.826 (02/99) - Error performance parameters and objectives for international, constant bit rate digital paths at or above the primary rate

[T1.103] ANSI T1.103 - 1987. Digital Hierarchy - Synchronous DS3 Format Specification

[T1.105] ANSI T1.105-1991. Digital Hierarchy - Optical Interface Rates and Format Specifications (SONET}

[T1.107] ANSI T1.107 - 1988. Digital Hierarchy - Format Specification

[T1.107a] ANSI T1.107a - 1990. Digital Hierarchy - Supplement to Format Specifications (DS3 Format Specifications)

[NANOG] St. Casner, C. Alaettinoglu, Ch. Kuan, A fine-grained view of high-performance networking, NANOG-22, May 2001

AUTHORS' ADDRESSES

Alexander ("Sasha") Vainshtein

Axerra Networks

24 Raoul Wallenberg St.

Tel Aviv 69719, Israel
email: sasha@axerra.com

Israel Sasson

Axerra Networks

24 Raoul Wallenberg St.

Tel Aviv 69719, Israel

email: israel@axerra.com

Akiva Sadovski

Axerra Networks

24 Raoul Wallenberg St.

Tel Aviv 69719, Israel

email: akiva@axerra.com

Eduard Metz

KPNQwest
Scorpius 60
2130 GE Hoofddorp, The Netherlands
email: eduard.metz@kpnqwest.com

Tim Frost

Zarlink Semiconductor

Tamerton Road, Roborough, Plymouth, PL6 7BQ, UK

email: tim.frost@zarlink.com

FULL COPYRIGHT STATEMENT

ACKNOWLEDGEMENT

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

ANNEX A. CESoPSN IN DIFFERENT TYPES OF PSN

A1. IP PSN

CESoPSN is RTP-based, and UDP flows are a natural way to convey RTP traffic (see [RFC1889]).
If this technique is used for conveying CESoPSN, then:
- o Unused even UDP ports must be allocated at both PE nodes terminating a CESoPSN PW as part of the PW establishment process

o IP and UDP headers must be prepended to each CESoPSN packet
o These packets will be transmitted by each PE node to its peer using the standard IP routing mechanisms.

UDP flows represent a multiplexing layer with limited ability to detect misconnections. As a consequence, SSRC-based misconnection detection by CESoPSN MAY be disabled.

IP represents a Carrier layer with inherent ability to infer the payload size from the header. As a consequence, detection of malformed packets SHOULD take the actual payload size into consideration.

By default, manual signaling can be used for setup and teardown of CESoPSN PWs over UDP flows. As a consequence, parameters defined in Section 6 should be incorporated into to the appropriate service-specific MIB module.

[RFC1889] defines a convention for associating an RTCP session with each RTP/UDP/IP one. Possible usage of RTCP for CESoPSN is left for further study.

A2. MPLS PSN

Note: The text below does not define a generic RTP/MPLS stack. Such a work is clearly out of scope of this document.

This section is concerned with the case of MPLS being used as both the PSN and multiplexing layer for the CESoPSN PW.

In this case, CESoPSN packet MUST be prepended with an MPLS label stack including:

o A VC label entry (see [MARTINI-TRANS] or [KOMPELLA]). This entry acts as the multiplexing layer header. It MUST be present in the stack and MUST be marked as residing at the bottom of the stack
o A tunnel label entry. This label, if present, acts as the PSN header and must immediately precede the VC label entry. It MAY be omitted in some situations.

This combination of PSN and multiplexing layers does not provide either frame length information or ability to detect misconnections. The former is not necessary for CESoPSN but limits ability to detect malformed packets in case of a very short packet payload. The

misconnection detection functionality can be provided using the following considerations:

1. The pattern in the first four bits following the bottom label ('1000') can be used as indication of an RTP header as it is distinct from any of the following:
   a. IPv4 pattern ('0100')
   b. IPv6 pattern ('0110')
   c. Pattern produced by Layer 2 services over MPLS encapsulated in accordance with [MARTINI-ENCAP] and using control word ('0000')
2. The SSRC field of the RTP header can be further used to detect misconnection.

MPLS tunnels are conventionally established using various signaling protocols. As a consequence, parameters used for setup and teardown of CESoPSN tunnels should be mapped to data elements of these protocols.

A3. L2TP PSN

Note: The text below does not define a generic RTP/L2TPv3 stack. Such a work is clearly out of scope of this document.

CESoPSN packets may be carried in L2TPv3 tunnels over IP (see [L2TPv3]) that would act as an alternative multiplexing layer over IP.

Since L2TPv3 provides both data and control plane for tunnel establishment, parameters describing payload and encapsulation layers should be defined as AVPs to allow single-ended setup and teardown of CESoPSN PWs.

L2TPv3 tunnels represent a multiplexing layer with an optional ability to detect misconnections using 32-bit or 64-bit "cookies". As a consequence, the PSN operator may choose between the L2TPv3-

based and SSRC-based misconnection detection techniques for CESoPSN PWs.

IP represents a PSN layer with inherent ability to infer the payload size from the header. As a consequence, malformed packets detection should consider actual payload size.

ANNEX B. EMULATION OF SONET/SDH CIRCUITS

B1. Relevant Types of SONET/SDH circuits

- o STS-1
- o STM-1

B2. Native Frame Size and Payload Format

Natural delineation of SONET/SDH frames (of abovementioned rates) will produce packets exceeding minimal MTU in some cases. As a consequence, a SONET/SDH frame must be fragmented into several CESoPSN packets will be used.

Usage of CESoPSN for unstructured SONET/SDH circuits requires presence of an appropriate framer in the ingress and egress PEs.

Each SONET/SDH frame will be fragmented into the Protocol Data Units (PDUs) of equal size. Data belonging to two and more different frames MUST NOT be combined into one PDU. For each SONET/SDH frame

only one CESoPSN packet will contain the framing octets (A1, A2) of this frame. Such a packet:

- o MUST contain these bytes aligned with its payload data (i.e., the 1st octet of the payload MUST contain the 1st A1 byte of a SONET/SDH frame
- o SHOULD be marked with M bit set to 1 in the RTP header.

B3. Synchronization modes

External clock sources traceable (in terms of G.781) to the same high quality (at least as defined in G.812) clock source should be available at both PEs for External or Differential timing.

B.3. Structure of the Control Word

The same bits as defined in Section 5.2.2 are used. However the meaning of the bits are slightly different:

- o Bit A - if set, represents LOS (e.g., as specified in [G.783]) of the incoming SONET/SDH signal. A packet with the A bit set should not carry any data
- o Bit I - if set, represents an Out-of-Frame (OOF) condition (e.g., as specified in [G.707]) of the incoming SONET/SDH signal. A packet with the I bit set should not carry any data

B4. Packetization and de-packetization

During normal operation, the CESoPSN packetizer will receive a fixed rate byte stream from a (physical or logical) SONET/SDH interface. When the whole SONET/SDH frame will be received, it will be partitioned into several blocks of equal size. After that, PSN and multiplexing headers are prepended to it and the resulting CESoPSN packets are transmitted into the PSN.
Because all normal CESoPSN packets associated with a specific SONET/SDH channel will have the same length, the transmission of CESoPSN packets for that channel SHOULD occur at regular intervals. At the far end of the packet network, the CESoPSN de-packetizer will receive packets into a jitter buffer, rebuild native SONET/SDH frames, and then play out the received byte stream at a fixed rate onto the corresponding PDH channel. The jitter buffer SHOULD be configurable to account for various network delay behavior patterns. The received 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 CESoPSN channel.
The RTP sequence numbers in the CESoPSN heard provide a mechanism to detect lost and/or reordered packets. The CESoPSN de-packetizer MUST detect lost or reordered packets.

B6. PSN to SONET/SDH Signals

Only CESoPSN defects requiring non-standard treatment are considered.

The CESoPSN de-packetizer MAY re-order packets received out of order. If the CESoPSN de-packetizer does not support re-ordering, it MUST drop out-of-order packets.

If any of the PDUs comprising a native SONET/SDH frame is lost, the scrambled pattern consisting of valid framing bytes ([G.707], [T1.105]) and all other bytes set to all 1s will be played out. The same pattern will be played out if a malformed packet has been detected.

The rationale for this behavior: an SDH node at the egress of a CESoPSN service may continue using the SDH signal received from the egress PE node as its clock source.