                        MAPOS/PPP Tunneling mode

Status of this Memo

Copyright Notice

IESG Note

   This memo documents a way of tunneling PPP over Sonet over MAPOS
   networks. This document is NOT the product of an IETF working group
   nor is it a standards track document. It has not necessarily
   benefited from the widespread and in depth community review that
   standards track documents receive.

Abstract

   This document specifies tunneling configuration over MAPOS (Multiple
   Access Protocol over SONET/SDH) networks. Using this mode, a MAPOS
   network can provide transparent point-to-point link for PPP over
   SONET/SDH (Packet over SONET/SDH, POS) without any additional
   overhead.

1. Introduction

   MAPOS [1][2] frame is designed to be similar to PPP over SONET/SDH
   (Packet over SONET/SDH, POS)[3][4] frame (Figure 1).

    a) MAPOS frame header (version 1)
       +----------+----------+----------+----------+
       | Address | Control | Protocol |
       | 8 bits | fixed,0x03| 16 bits |
       +----------+----------+----------+----------+

    b) MAPOS frame header (MAPOS 16)
       +----------+----------+----------+----------+
       | Address | Protocol |
       | 16bits | 16 bits |
       +----------+----------+----------+----------+

    c) PPP frame header
       +----------+----------+----------+----------+
       | Address | Control | Protocol |
       | fixed,0xFF| fixed,0x03| 16 bits |
       +----------+----------+----------+----------+

    Figure 1. Header similarity of MAPOS frame and POS frame

This means that a MAPOS network can easily carry POS frames with no
additional header overhead by rewriting only 1 or 2 octets. PPP
tunneling configuration over MAPOS networks (MAPOS/PPP tunneling
mode) provides for efficient L2 multiplexing by which users can share
the cost of high speed long-haul links.

This document specifies MAPOS/PPP tunneling mode. In this mode, a
MAPOS network provides a point-to-point link for those who intend to
connect POS equipment. Such link is established within a MAPOS
switch, or between a pair of MAPOS switches that converts between POS
header and MAPOS header for each L2 frame.

Chapter 2 describes the specification in two parts. First part is
user network interface (UNI) specification and the second part is
operation, administration, management and provisioning (OAM&P)
description. Other issues such as congestion avoidance, end-to-end
fairness control are out of scope of this document.

Implementation issues are discussed in Chapter 3. Security
considerations are noted in Chapter 4.

2. MAPOS/PPP tunneling mode

2.1 Overview

   MAPOS/PPP tunneling mode is based on header rewriting. Figure 2.
   shows an example of MAPOS/PPP tunneling mode. The MAPOS network uses
   MAPOS 16 [2] in this example. Consider a tunneling path between
   customer premise equipment (CPE) A and CPE B which are industry
   standard POS equipment. The ingress/egress MAPOS switches A/B
   assigns unique MAPOS addresses (0x0203 and 0x0403) to the CPEs.
   These MAPOS addresses are used in the MAPOS network, for frame
   forwarding between CPE A and CPE B. NSP [5] will not be running
   between the CPEs and the switches in this case.

   MAPOS switch A rewrites the first 2 octets of every frame from CPE A,
   which are fixed as 0xFF and 0x03, to the MAPOS address of its peer,
   which is 0x0403. Frames are forwarded by the MAPOS network and
   arrives at the egress MAPOS switch B which rewrites the first 2
   octets to their original values. If MAPOS v1 [1] is used in the
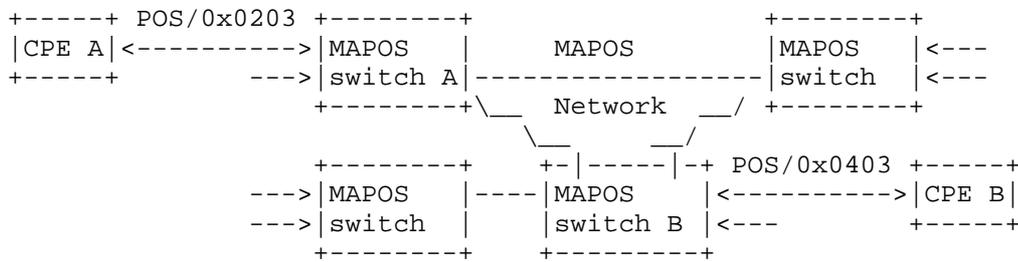   MAPOS network, only the first octet is rewritten.

```
   +-----+ POS/0x0203 +--------+                   +--------+
   |CPE A|<---------->|MAPOS   |     MAPOS         |MAPOS   |<---
   +-----+       --->|switch A|-------------------|switch  |<---
                     +--------+\__  Network  __/ +--------+
                                \__        __/
                     +--------+    +-|-----|-+ POS/0x0403 +-----+
                 --->|MAPOS   |----|MAPOS   |<---------->|CPE B|
                 --->|switch  |    |switch B|<---          +-----+
                     +--------+    +---------+
```

                  Figure 2. MAPOS/PPP tunneling mode

   The tunneling path between the two CPEs is managed by the
   ingress/egress MAPOS switches.

2.2 User-Network Interface(UNI)

2.2.1 Physical Layer

   For transport media between border MAPOS switch and CPE, SONET/SDH
   signal is utilized. Signal speed, path signal label, light power
   budget and all physical requirements are the same as those of PPP
   over SONET/SDH [3].

SONET/SDH overheads are terminated at the ingress/egress switches.
SONET/SDH performance monitors and alarms are used for the link
management between a CPE and the switch. Inter-switch links are
similarly managed by SONET/SDH monitors and alarms.

A CPE should synchronize to the clock of the border MAPOS switch.
The corresponding port of the MAPOS switch uses its internal clock.
When the CPE is connected to the MAPOS switch through SONET/SDH
transmission equipment, both should synchronize to the clock of the
SONET/SDH transmission equipment.

2.2.2 Link layer

Link layer framing between CPE and MAPOS switch also follows the
specification of PPP over SONET/SDH [3].

HDLC operation including byte stuffing, scrambling, FCS generation is
terminated at the ingress/egress switch. In a MAPOS switch, HDLC
frame [4] is picked up from a SONET/SDH payload and the first octet
(HDLC address) for MAPOS v1 [1], or the first two octets (HDLC
address and control field) for MAPOS 16 [2] are rewritten. The
operation inside the border switch is as follows:

 From CPE (Ingress Switch receiving):

    SONET/SDH framing
       -> X^43+1 De-scrambling -> HDLC Byte de-stuffing
       -> HDLC FCS detection (if error, silently discard)
       -> L2 HDLC address/control rewriting
           (0xFF   -> MAPOS v1 destination address, or
            0xFF03 -> MAPOS 16 destination address)
       -> MAPOS-FCS generation
       -> HDLC Byte stuffing -> X^43+1 Scrambling -> SONET/SDH framing

  To CPE (Egress Switch transmitting):

    SONET/SDH framing
       -> X^43+1 De-scrambling -> HDLC Byte de-stuffing
       -> MAPOS-FCS detection (if error, silently discard)
       -> L2 HDLC address/control rewriting
           (MAPOS v1 address -> 0xFF, or
            MAPOS 16 address -> 0xFF03)
       -> HDLC FCS generation
       -> HDLC Byte stuffing -> X^43+1 Scrambling -> SONET/SDH framing

For STS-3c-SPE/VC-4, non-scrambled frame can be used for
compatibility with RFC 1619. However, the use of 32bit-CRC and
X^43+1 scrambling is recommended in RFC2615 [3] and for MAPOS
networks.

Maximum transmission unit (MTU) of the link must not be negotiated
larger than MAPOS-MTU which is 65280 octets.

Figure 3 shows a CPE-side L2 frame and the converted frame in the
ingress/egress MAPOS switches. Note that the MAPOS/PPP tunneling
mode is not a piggy-back encapsulation, but it is a transparent link
with no additional header overhead.

```
<--- Transmission
      +----------+----------+----------+----------+
      |   Flag   | Address  | Control  | Protocol |
      | 01111110 | 11111111 | 00000011 | 16 bits  |
      +----------+----------+----------+----------+
      +------------+---------+----------+----------+----------------
      | Information | Padding |HDLC FCS  |   Flag   | Inter-frame Fill
      |     *       |    *    |16/32 bits| 01111110 | or next Address
      +------------+---------+----------+----------+----------------

        (a) HDLC frame from/to CPE

<--- Transmission
      +----------+----------+----------+----------+
      |   Flag   | MAPOS Destination   | Protocol |
      | 01111110 | 0xxxxxx0 | xxxxxxx1 | 16 bits  |
      +----------+----------+----------+----------+
      +------------+---------+----------+----------+----------------
      | Information | Padding |MAPOS FCS |   Flag   | Inter-frame Fill
      |     *       |    *    |16/32 bits| 01111110 | or next Address
      +------------+---------+----------+----------+----------------

        (b) Converted MAPOS 16 frame, forwarded in MAPOS networks
```

       Figure 3. HDLC frame from/to CPE and its conversion

2.3 Operation, Administration, Management and Provisioning (OAM&P)

2.3.1 MAPOS/PPP mode transition

When a port of MAPOS switch is configured to PPP tunneling mode, at
least the following operations are performed in the switch.

    a) Disable NSP [5] and SSP [6] (for the port, same below)
    b) Disable MAPOS broadcast and multicast forwarding

       c) Reset the Path Signal Label (C2) to 0x16 if X^43+1 scrambling
          is used. The value 0xCF is used for non-scrambled OC3c signal.
       d) Enable header rewriting function to specified destination
          address

   When the port is configured back to MAPOS mode, reverse order of the
   operations above are performed. That means;

       a) Disable header rewriting function (for the port, same below)
       b) Reset the Path Signal Label (C2) to MAPOS default (0x8d)
       c) Enable MAPOS broadcast and multicast forwarding
       d) Enable NSP and SSP

   SONET/SDH alarms (B1/B2/B3 error exceeding, SLOF, SLOS, etc.) should
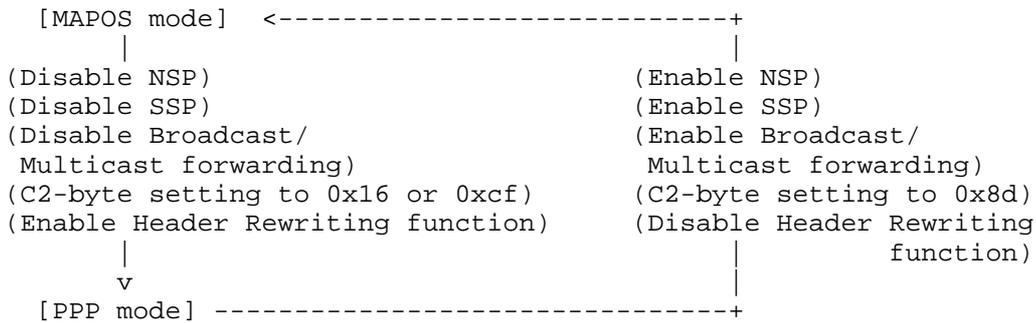   not affect this transition. Figure 4 shows mode transition described
   above.

```
     [MAPOS mode]  <----------------------------+
         |                                       |
 (Disable NSP)                           (Enable NSP)
 (Disable SSP)                           (Enable SSP)
 (Disable Broadcast/                     (Enable Broadcast/
  Multicast forwarding)                   Multicast forwarding)
 (C2-byte setting to 0x16 or 0xcf)      (C2-byte setting to 0x8d)
 (Enable Header Rewriting function)     (Disable Header Rewriting
         |                                   |            function)
         v                                   |
     [PPP mode] -----------------------------+
```

        Figure 4. MAPOS/PPP tuneling mode state transition diagram

2.3.2 Path Establishment

   A MAPOS/PPP tunneling path is established by following steps.

       a) Choose MAPOS address pair on both ingress/egress switches and
          configure their ports to PPP tunneling mode (see 2.3.1).

       b) When the routes for both directions become stable, the
          tunneling path is established. The link between the CPEs may
          be set up at that moment; PPP LCP controls are transparently
          exchanged by the CPEs.

   To add a new path, operators should pick unused MAPOS address-pair.
   They may be determined simply by choosing switches and ports for each
   CPE, because there is one-to-one correspondence between MAPOS
   addresses and switch ports.

Then, those ports should be configured to MAPOS/PPP tunneling mode on
both of the switches. Frame reachability is provided by SSP [6] in
the MAPOS network. When the frame forwarding for each direction are
stable, the path is established and frame forwarding is started.
Until then, the link between border switches and CPE should be down.

A MAPOS/PPP tunneling path should be managed by the pair of MAPOS
addresses. It should be carefully handled to avoid misconfiguration
such as path duplication. For convenient management, path database
can be used to keep information about pairs of MAPOS addresses. Note
that the path database is not used for frame forwarding. It is for
OAM&P use only.

2.3.3 Failure detection and indication

When any link or node failure is detected, it should be indicated to
each peer of the path. This is done by PPP [7] keep-alive (LCP Echo
request/reply) for end-to-end detection.

Consideration is required to handle SONET/SDH alarms. When a link
between CPE and the MAPOS switch fails, it is detected by both the
MAPOS switch and the CPE seeing SONET/SDH alarms. However, far-side
link remains up and no SONET/SDH error is found; SONET/SDH alarms
are not transferred to the far end because each optical path is
terminated in MAPOS network. In this case, the far end will see
'link up, line protocol down' status due to keep-alive expiration.

For example, Figure 5 shows a tunneling path. When link 1 goes down,
MAPOS sw A and CPE A detects SONET/SDH alarms but MAPOS sw B and CPE
A' do not see this failure. When PPP keep-alive expires, CPE A'
detects the failure and stops the packet transmission. The same
mechanism is used for failure within the MAPOS cloud (link 2). When
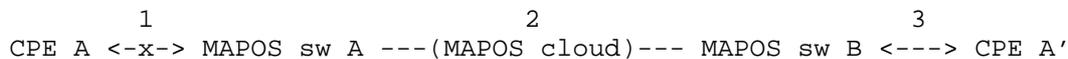a MAPOS switch is down, SSP handles it as a topology change.

```
                1                      2                      3
     CPE A <-x-> MAPOS sw A ---(MAPOS cloud)--- MAPOS sw B <---> CPE A'
```

                        Figure 5. Link failure

2.3.4 Path removal

A MAPOS/PPP tunneling path is removed by following steps.

    a) Choose the path to remove, configure MAPOS switches on both
       ends of the path to disable the ports connected to the CPEs.

    b) Path database may be updated that the path is removed.

        c) When CPE is detached, port may be reset to MAPOS default
           configurations.

   Frames arriving after the destination port was disabled should be
   silently discarded and should not be forwarded to the port.

2.3.5 Provisioning and Design Consideration

   Because MAPOS does not have any QoS control at its protocol level,
   and POS does not have flow-control feature, it is difficult to
   guarantee end-end throughput. Sufficient bandwidth for inter-switch
   link should be prepared to support all paths on the link.

   Switches are recommended to ensure per-port fairness using any
   appropriate queuing algorithm. This is especially important for
   over-subscribed configuration, for example to have more than 16 OC12c
   paths on one OC192c inter-switch link.

   Although MAPOS v1 can be applied to the MAPOS/PPP tunneling mode,
   MAPOS 16 is recommended for ease of address management.

   Automatic switch address negotiation mechanism is not suitable for
   the MAPOS/PPP tunneling mode, because the path management mechanism
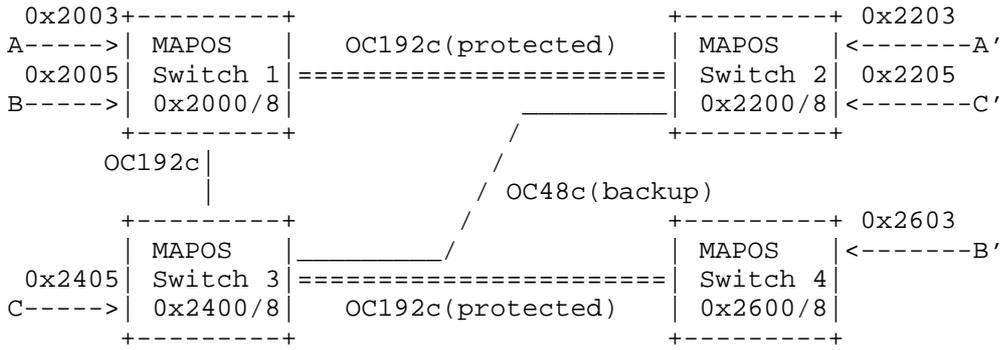   becomes much more complex.

3. Implementation

3.1 Service example

   Figure 6 shows an example of MAPOS network with four switches.
   Inter-switch links are provided at OC192c and OC48c rate, customer
   links are either OC3c or OC12c rate. Some links are optically
   protected. Path database is used for path management.

   Using MAPOS-netmask with 8 bits, this topology can be extended up to
   64 MAPOS switches, each equipped with up to 127 CPE ports. Switch
   addresses are fixed to pre-assigned values.

   The cost of optical protection (< 50ms) can be shared among paths.
   Unprotected link can also be coupled for more redundancy in case of
   link failure. SSP provides restoration path within few seconds.

```
  0x2003+---------+                          +---------+ 0x2203
   A----->| MAPOS   |    OC192c(protected)    | MAPOS   |<-------A'
   0x2005| Switch 1|=========================| Switch 2| 0x2205
   B----->| 0x2000/8|          _____     | 0x2200/8|<-------C'
         +---------+         /               +---------+
     OC192c|                /
          |              / OC48c(backup)
         +---------+     /                     +---------+ 0x2603
         | MAPOS   |_____/                 | MAPOS   |<-------B'
   0x2405| Switch 3|=========================| Switch 4|
   C----->| 0x2400/8|   OC192c(protected)     | 0x2600/8|
         +---------+                           +---------+


     Path database entries:
     -------------------------------------------------------------
      User : Speed : Mode               : Address pair  : Status
     -------------------------------------------------------------
      A-A' : OC3c  : CRC32, scramble : 0x2003-0x2203 : Up and running
      B-B' : OC12c : CRC32, scramble : 0x2005-0x2603 : B Down
      C-C' : OC3c  : CRC16, no-scram : 0x2405-0x2205 : C' Down
     -------------------------------------------------------------
```

         Figure 6. Example Topology and its Path Management

3.2 Evaluation of latency of reference implementation

   Figure 7 shows evaluation platforms in terms of latency measurement
   of MAPOS/PPP tunneling mode.

Case 1: Base latency measurement

```
     Measurement
     Equipment
     +---------+   POS Unidirectional Flow, OC12c 30%, FCS 32bits,
     | IXIA 400|   payload-scrambling on (same for all cases)
     | POS-LM  |<--+
     | OC12c x2|---+ Loopback
     +---------+
     (Using IxSoftware v3.1.148/SP1d)
```

Case 2: Router latency measurement

```
     Measurement                          Device Under Test
     +---------+  POS                     +------------+
     | IXIA 400|  Unidirectional Flow | Cisco GSR  |
     | POS-LM  |<--------------------| 12008/1port|
     | OC12c x2|-------------------->| OC12cLC x2 |
     +---------+                         +------------+
                                  (Using IOS 12.0(15)S1)
```

Case 3: MAPOS/PPP tunneling switch latency measurement

```
     Measurement                          Device Under Test
     +---------+  POS                     +-------------+
     | IXIA 400|  Unidirectional Flow | CSR MAPOS    |
     | POS-LM  |<--------------------| CORESwitch80|
     | OC12c x2|-------------------->| OC12c x2     |
     +---------+                         +-------------+
```

Figure 7. Latency measurement of reference platform for MAPOS/PPP
tunneling mode

There is a PPP connection between port 1 and 2 of the measurement
equipment. Traffic comes from measurement equipment (IXIA 400) and
forwarded by a device under test back to the equipment. Timestamping
and latency calculation are performed by IXIA 400 automatically.
Traffic Load is set to 30% of OC12c for offloading router.

Results are shown in Table 1. Measurements were taken according to
the RFC2544 requirements [8]. We measured 25 trials of 150 seconds
duration for each frame size. Results are averaged and rounded to
the 20 ns resolution of IXIA. 95% confidence interval (C.I.) value
are also rounded.

```
     -----------------------------------------------------------------
     Frame size (bytes)   64    128    256    512   1024   1280   1518
     -----------------------------------------------------------------
     Latency(ns)
     -----------------------------------------------------------------
     Case 1: Baseline   4060   5640   6940   9840  16420  20700  23340
        95% C.I.(+/-)     20     80     60    180     80    100    120
     -----------------------------------------------------------------
     Case 2: Router    26560  28760  33860  44600  68280  80500  91160
        95% C.I.(+/-)    200    100    160    220    100    100    200
     -----------------------------------------------------------------
     Case 3: Switch    11100  13480  16620  22920  36380  43900  49920
        95% C.I.(+/-)    120    120    120    200    100    160    120
     -----------------------------------------------------------------
```
              Table 1. Results of Latency (ns) - Frame size (bytes)

   This results shows that MAPOS/PPP tunneling mode does not cause any
   performance degradation in terms of latency view. A POS L2 switch
   was reasonably faster than a L3 router.

4. Security Considerations

   There is no way to control or attack a MAPOS network from CPE side
   under PPP tunneling mode. It is quite difficult to inject other
   stream because it is completely transparent from the viewpoint of the
   CPE. However, operators must carefully avoid misconfiguration such
   as path duplication. Per-path isolation is extremely important;
   switches are recommended to implement this feature (like VLAN
   mechanism).

   In addition, potential vulnerability still exists in a mixed
   environment where PPP tunneling mode and MAPOS native mode coexists
   in the same network. Use of such environment is not recommended,
   until an isolation feature is implemented in all MAPOS switches in
   the network. Note that there is no source address field in the MAPOS
   framing, which may make path isolation difficult in a mixed MAPOS/PPP
   environment.

5. References

   [1]   Murakami, K. and M. Maruyama, "MAPOS - Multiple Access Protocol
           over SONET/SDH Version 1", RFC 2171, June 1997.

   [2]   Murakami, K. and M. Maruyama, "MAPOS 16 - Multiple Access
           Protocol over SONET/SDH with 16 Bit Addressing", RFC 2175, June
           1997.

   [3]   Malis, A. and W. Simpson, "PPP over SONET/SDH", RFC 2615, June
           1999.

   [4]   Simpson, W., "PPP in HDLC-like Framing", STD 51, RFC 1662, July
           1994.

   [5]   Murakami, K. and M. Maruyama, "A MAPOS version 1 Extension -
           Node Switch Protocol," RFC 2173, June 1997.

   [6]   Murakami, K. and M. Maruyama, "A MAPOS version 1 Extension -
           Switch-Switch Protocol", RFC 2174, June 1997.

   [7]   Simpson, W., "The Point-to-Point Protocol (PPP)", STD 51, RFC
           1661, July 1994.

   [8]   Bradner, S. and J. McQuaid, "Benchmarking Methodology for
           Network Interconnect Devices", RFC 2544, March 1999.

6. Acknowledgments

   The authors would like to acknowledge the contributions and
   thoughtful suggestions of Takahiro Sajima.

7. Author's Address

   Susumu Shimizu
   NTT Network Innovation Laboratories,
   3-9-11, Midori-cho Musashino-shi
   Tokyo 180-8585 Japan

   Phone: +81 422 59 3323
   Fax: +81 422 59 3765
   EMail: shimizu@ntt-20.ecl.net

   Tetsuo Kawano
   NTT Network Innovation Laboratories,
   3-9-11, Midori-cho Musashino-shi
   Tokyo 180-8585 Japan

   Phone: +81 422 59 7145
   Fax: +81 422 59 4584
   EMail: kawano@core.ecl.net

   Ken Murakami
   NTT Network Innovation Laboratories,
   3-9-11, Midori-cho Musashino-shi
   Tokyo 180-8585 Japan

   Phone: +81 422 59 4650
   Fax: +81 422 59 3765
   EMail: murakami@ntt-20.ecl.net

   Eduard Beier
   DeTeSystem GmbH
   Merianstrasse 32
   D-90409 Nuremberg, Germany

   EMail: Beier@bina.de

8. Full Copyright Statement

Acknowledgement