PTP over MPLS
draft-ronc-ptp-mpls-00

Status of this Memo

By submitting this Internet-Draft, each author represents that any applicable patent or other IPR claims of which he or she is aware have been or will be disclosed, and any of which he or she becomes aware will be disclosed, in accordance with Section 6 of BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

The list of current Internet-Drafts can be accessed at http://www.ietf.org/ietf/1id-abstracts.txt.

The list of Internet-Draft Shadow Directories can be accessed at http://www.ietf.org/shadow.html.

This Internet-Draft will expire on January 1, 2008.

Copyright Notice

Copyright (C) The IETF Trust (2007).

Abstract

This document defines mapping of the Precision Time Protocol for version 2 and higher over MPLS.
### Table of Contents

1. Requirements Language ............................................. 3  
2. Acronyms ..................................................................... 3  
3. Introduction .................................................................... 3  
4. Encapsulation ................................................................... 5  
5. PTP LSPs ....................................................................... 6  
6. MPLS PW Setup and Signaling .......................................... 9  
   6.1. General ..................................................................... 9  
   6.2. PW Type .................................................................... 9  
   6.3. FEC .......................................................................... 10  
7. IANA Considerations ...................................................... 12  
8. Security Considerations .................................................. 12  
9. References ...................................................................... 12  
   9.1. Normative References ............................................... 12  
   9.2. Informative References ............................................. 13  
Author’s Address ............................................................... 13  
Intellectual Property and Copyright Statements ..................... 14
1. Requirements Language

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].

2. Acronyms

- AII: Attachment Individual Identifier
- AGI: Attachment Group Identifier
- FEC: Forwarding Equivalence Class
- GID: Generalized pseudowire IDentifier
- P2P: Peer to Peer
- E2E: End to End
- LDP: Label Distribution Protocol
- LSP: Label Switched Path
- LSR: Label Switch Router
- MPLS: Multi Protocol Label Switching
- PE: Provider Edge
- PTP: Precision Time Protocol (IEEE1588)
- PW: Pseudowire
- RSVP-TE: Resource Reservation Protocol-Traffic Engineering
- VCCV: Virtual Circuit Connection Verification

3. Introduction

The Precision Time Protocol (PTP) objective is to synchronize independent clocks running on separate nodes of a distributed system. PTP, aka IEEE1588, was originally developed for the measurement and control industry. The second version of the standard added provisions for higher accuracy, higher and varied update rates, introduced the concept of transparent clocks and unicast
communication between clocks. PTPv2 modified the message formats and standardized the transport of PTP over UDP/IPv6 and directly over Ethernet in addition to the UDP/IPv4 defined in version 1. This document defines direct mapping of PTP over MPLS.

PTP defines intermediate clock functions between the source of time (Grandmaster) and the slave clocks, called boundary clocks and transparent clocks. Boundary clocks form master-slave hierarchy with the Grandmaster clock as root. A boundary clock has several ports; one PTP port is selected as slave port, while the other ports become either master ports or passive ports according to the master-slave hierarchy. The messages related to synchronization, establishing the master-slave hierarchy, and signaling, terminate in the protocol engine of a boundary clock and are not forwarded. Management messages are forwarded to other ports on the boundary clock subject to restrictions to limit the propagation of these messages within the system. Two types of transparent clocks are defined, end-to-end (E2E) and peer-to-peer (P2P). Transparent clocks modify a 'correction field' within the synchronization messages to compensate for residence and propagation delays. Transparent clocks do not terminate synchronization, master-slave hierarchy control messages or signaling messages.

PTP messages over Ethernet or IP can always be tunneled over MPLS. The direct PTP over MPLS mapping is applicable whenever either the edge or the core of the MPLS network is PTP aware. For example, edge Label Switch Router (E-LSR) that acts as PTP boundary clock may use PTP over MPLS encapsulation towards the MPLS cloud while it may use other PTP encapsulations outside of the MPLS cloud. Boundary clock function may be required to enhance performance, to adapt between networks or for administrative reasons. Direct PTP over MPLS mapping is more bandwidth efficient than any non-direct mapping. Direct MPLS mapping allows easier introduction of PTP transparent clocks within the MPLS cloud, as it provides easier identification of PTP event messages compared with other native payloads and ensures that the correction field is located close enough to the MPLS label for efficient hardware implementation. PTP over MPLS follows the mapping of pseudowires over MPLS and benefits from the umbrella of protocols for discovery, setup, negotiation and maintenance protocols associated with it.

PTP supports both a multicast and a unicast communication model. Within the multicast communication model clocks automatically discover other clocks within their communication path and reorganize themselves in a master slave hierarchy. Synchronization and master-slave hierarchy messages are sent in multicast by the master to all slaves within the communication path. In the unicast communication model, clocks are pre-configured with the addresses of the other
clocks within their communication path, and all communication is performed in unicast. A combined model is also possible, e.g. in which master-slave hierarchy messages are sent in multicast, while synchronization messages are sent in unicast. MPLS is a connection oriented technology. MPLS label switched paths (LSPs) must be set up in advance in order to deliver messages between nodes. MPLS supports point-to-point as well as point-to-multipoint and multipoint-to-multipoint LSPs. However this document describes the mapping of PTP over MPLS using point to point MPLS LSPs only, following the PTP unicast communication model. Use of point-to-multipoint or multipoint-to-multipoint LSPs are left for further study.

4. Encapsulation

This section defines the PTP over MPLS encapsulation format.

```
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|            Tunnel Label               | Exp |S|       TTL     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                               PW label                | Exp |S|       TTL     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|               TS |             PTP Header (Cont.)                        |
|0 0 0 0|                                                       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     PTP Header (Cont.)                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     PTP Payload                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 1: PTP over MPLS encapsulation

Figure 1 describes PTP over MPLS mapping. The PTP header defined in [PTPv2] immediately follows the bottom MPLS label. The first nibble of the PTP header is called transportSpecific (TS) and is defined as follows: "The transportSpecific field may be used by a lower layer transport protocol and is defined by the mapping specification of that protocol in an annex of this standard". Non-zero values of the transportSpecific fields are used for various purposes in PTP over UDP/IPv4 (PTPv1 backward compatibility) and PTP over Ethernet.
(subtype) mappings. The transportSpecific field MUST be set to zero on transmit of all PTP over MPLS messages. Only messages received with transportSpecific field set to zero MUST be processed as valid PTP over MPLS messages.

The PTP over MPLS mapping is consistent with the generic PW MPLS control word structure defined in [RFC4385]. [RFC4385] specifies the preferred design of a Pseudowire Emulation Edge-to-Edge (PWE3) Control Word to be used over an MPLS packet switched network, and the Pseudowire Associated Channel Header. The design of these fields is chosen so that an MPLS Label Switching Router performing MPLS payload inspection will not confuse a PWE3 payload with an IP payload. The first nibble immediately following the bottom MPLS label is 0000b, while this nibble is either 0100b for IPv4 over MPLS and 0110b for IPv6 over MPLS.

The Psuedowire Associated Channel Header (PACW) is defined in [RFC4385]. The value of the first nibble following the bottom MPLS label is 0001b. An associated channel is a channel that is multiplexed in the PW with PTP traffic, and thus follows the same path through the PSN as PTP does. The associated channel can be used as an Operation, Administration and Management (OA&M) channel. VCCV [VCCV] defines Virtual Circuit Connection Verification (VCCV) protocol that can be run over the associated channel. VCCV includes bi-directional connectivity checks for fault detection and fault and status signaling. Use of VCCV for PTP over MPLS pseudowires is for further study.

The PW architecture [RFC3985] specifies the use of a label stack with at least 2 levels. The label at the bottom of the stack is called the PW label. The PW label acts as an identifier for a particular PW. Labels higher in the stack are called the packet switch network (PSN) labels, or tunnel labels, and are used to forward the packet through the MPLS network as described in [RFC3031]. Use of the PW label within PTP is detailed in the next section.

5. PTP LSPs

PTP defines ports, communication paths and domains. A PTP port is logical access point of a PTP clock for PTP communications to the communications network. A communication path is a segment of a network enabling direct communication between two PTP ordinary or boundary clocks. A PTP domain is a logical grouping of PTP clocks that synchronize to each other using the PTP protocol, but that are not necessarily synchronized to PTP clocks in another PTP domain.

Over MPLS a communication path is a set of Label Switched Paths
(LSPs). There are two layers of LSPs, the MPLS tunnel LSP that switches the MPLS tunnel label(s) and the PTP (PW) LSP that switches the PW label. The tunnel LSP connects two PTP-aware LSRs. For example, the tunnel LSP may span multiple LSRs connecting two ordinary or boundary clocks each residing on an opposite edge of an MPLS network. The PTP LSP connects two PTP aware-LSRs and is switched only by LSRs that are also PTP transparent clocks.

Figure 2: Tunnel and PTP LSPs

Figure 2 is an example of tunnel and PTP LSPs formed between a boundary clock, an ordinary clock and an E2E transparent clock LSRs. The boundary clock is the Edge-LSR of a tunnel LSP connecting it with the transparent clock LSR. The Transparent clock LSR has another tunnel LSP connecting it with the Ordinary Clock LSR. A single PTP LSP connects the boundary clock to the ordinary clock. The E2E transparent clock LSR corrects the residence time of PTP event messages switched over the PTP LSP. A PTP LSP that connects between two ports of ordinary or boundary clocks is called E2E PTP LSP.

A P2P transparent clock requires exchange of peer-to-peer delay messages with its peers. Assume the transparent clock of Figure 2 is P2P transparent clock, then in addition to the PTP LSP between the boundary clock and the ordinary clock, two LSPs between the P2P transparent clock and its peers, i.e. between the transparent clock and the ordinary and between the transparent clock and the boundary clock, must be formed as well. These LSPs will carry peer-to-peer delay messages for measuring the delay between the peers. A PTP LSP that connects between two peer ports for exchange of peer-to-peer delay messages is called a PTP P2P LSP. PTP peer-to-peer messages MUST only be sent on PTP P2P LSPs. Other PTP message types MUST NOT be sent over PTP P2P LSPs.

PTP Management messages are sent between PTP clocks and management
nodes. PTP management messages can be sent to a particular port on a particular clock, to the clock itself or to all clocks within the PTP domain. The latter PTP messages should be relayed by boundary clocks to all PTP ports within the communication path. PTP management messages include time-to-live counters to avoid forwarding loop. PTP management messages are not required if other means of configuration and monitoring is used within the domain. It is most likely that within MPLS networks PTP clocks will be managed and monitored by SNMP or similar protocols, and PTP management messages will not be used. The mapping of PTP management messages over MPLS is left for future versions of this document.

Within a PTP communication path, all PTP ports of ordinary and boundary clocks must be able to communicate with all other potential master ports of the communication path. Slave only clock ports communicate only with the master clock of the communication path (and possibly with management nodes). Therefore each PTP port of ordinary or boundary clock that belongs to a communication path MUST form a separate single PTP E2E LSP with each other PTP non-slave-only port of an ordinary or boundary clock that belongs to the same communication path. In particular, the master port of the communication path must send separate Sync message on each of the E2E PTP LSPs connecting to each of the other ordinary or boundary ports within the communication path. If the communication path supports the peer-to-peer delay mechanism, all PTP ports MUST establish in addition a PTP P2P LSP with their peers.

When more than one PTP domain is used within an MPLS cloud, separate P2P and E2E LSPs MUST be established to connect ports of clocks belonging to different domains. When applicable, the same tunnel LSP can be used to deliver PTP LSPs of more than one domain.

More than one PTP port MAY be tunneled over the same MPLS tunnel LSP. Different PTP LSPs MUST be used to transport PTP messages sent from different ports.

All PTP LSPs MUST be bi-directional. PTP protocol is sensitive to asymmetry between the propagation delays across the master-slave and slave-master paths. If the asymmetry is known, the PTP protocol has means to compensate for the known asymmetry. If the asymmetry is not known, the time recovered at the slaves will have a fixed offset equal to half the propagation delay asymmetry. Therefore, whenever possible, the tunnel and PTP LSPs SHOULD be setup in a symmetric way.

A P2P transparent clock corrects PTP event messages for residence and upstream propagation delay. The upstream propagation delay to its peer is calculated in advance using the peer-to-peer delay messages. An MPLS P2P transparent clock associates the peer delay computed over
the PTP P2P LSP with the E2E LSP sent from the same PTP port and domain, over the same MPLS tunnel. Therefore in a P2P communication path each PTP port MUST be associated with a single MPLS tunnel LSP. If connectivity requires multiple MPLS tunnels, each MPLS tunnel is considered as a separate communication path, and MUST be associated with a different PTP port.

6. MPLS PW Setup and Signaling

6.1. General

PTP PWs MUST be set up in advance for the transport of PTP messages. Any acceptable method of MPLS label distribution MAY be used for distributing the MPLS tunnel label [RFC3031]. These methods include LDP [RFC3036], RSVP-TE [RFC3209], or configuration. Any acceptable method of PW label distribution MAY be used for distributing the PW label. This section discusses the distribution of PW labels using the LDP protocol. [RFC4447] specifies the MPLS label distribution protocol (LDP) [RFC3036] extensions to set up and maintain the PWs, and defines new LDP objects to identify and signal attributes of PWs.

To assign and distribute the PW labels, an LDP session MUST be set up between the PW endpoints using the extended discovery mechanism described in [RFC3036]. The PW label bindings are distributed using the LDP downstream unsolicited mode described in [RFC3036]. An LDP label mapping message contains a FEC object, a label object, and possible other optional objects. The FEC object indicates the meaning of the label, identifies the PW type, and identifies the PW that the PW label is bound to. See [RFC4447] for further explanation of PW signaling.

6.2. PW Type

This specification defines new PW type values to be carried within the FEC object to identify PTP PWs. The PW type is a 15-bit parameter assigned by IANA, as specified in the [RFC4446] registry, and MUST be used to indicate the PTP LSP type being used on the PW. Two PW types are used for PTP PWs:

<table>
<thead>
<tr>
<th>PW type Description</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>TBD A E2E PTP: End-to-end PTP PW</td>
<td></td>
</tr>
<tr>
<td>TBD B P2P PTP: Peer-to-peer PTP PW</td>
<td></td>
</tr>
</tbody>
</table>

PTP PWs MUST be bi-directional, which means that a unidirectional leg of the PW MUST be set up in each direction. The same PW type MUST be
used for PW signaling in both directions.

6.3. FEC

PTP PWs MUST use the Generalized PWid (GID) FEC element as defined in [RFC4447]. The GID FEC element includes TLV fields for attachment individual identifiers (AII) that, in conjunction with an attachment group identifier (AGI), serve as PW endpoint identifiers. The endpoint identifier on the local PE (denoted as <AGI, source AII or SAII>) is called the source attachment identifier (SAI) and the endpoint identifier on the remote PE (denoted as <AGI, target AII or TAII>) is called the target attachment identifier (TAI).

The PTP port identity and the PTP domain number form a unique PTP PW endpoint identifier. The PTP domain number is common to both endpoints and can therefore serve as the AGI. The PTP port identity can be used as AII. The PTP port identity is composed of an 8-octet clock identity and a 16-bit unsigned integer port number. The clock identity MUST be either globally unique or unique within the domain. Two formats are defined in [PTPv2]; the preferred clock identity is formatted as IEEE EUI-64 address. IEEE MAC-48 address can also be used as EUI-64 globally unique clock identity. The second possible clock identity format is defined for technology specific or closed systems. An MPLS specific clock identity format that follows the aggregation friendly AII type [AGGAII] may be considered in future versions of this standard.
Figure 3: PTP Generalized PWid FEC element

Figure 3 describes the GID FEC used for PTP PW. The PTP Type is set to either P2P or E2E PTP PW type. A new AGI and AII types MUST be used, with values to be assigned by IANA:

AII type Description

TBD C PTP PortId: PTP Port Identity

AGI type Description

TBD D PTP Domain: PTP domain number

The PTP domain number is mapped to the first octet immediately following the length octet of the AGI sub TLV. The PTP source port number immediately follows the length octet of the SAII sub TLV followed by 8 octets of the source clock identity. The target port identity is encoded similarly.
Use of the Interface Parameters TLV and the PW Grouping TLV defined in [RFC4447] for PTP PW is left for future versions of this document.

7. IANA Considerations

This document requests that IANA allocate a value from the "Attachment Individual Identifier (AII) Type" registry defined in [RFC4446] as PTP PortId AII Type.

The suggested value for this AII type is 0x03.

This document requests that IANA allocate a value from the "Attachment Group Identifier (AGI) Type" registry defined in [RFC4446] as PTP Domain AGI Type.

The suggested value for this AGI type is 0x02.

This document requests that IANA allocate two values from the "MPLS Pseudowire Type" registry defined in [RFC4446] as E2E PTP PW Type and P2P PW Type.

The suggested value for the E2E PTP type is 0x001F. The suggested value for the P2P PTP type is 0x0020.

8. Security Considerations

MPLS PW security considerations in general are discussed in [RFC3985] and [RFC4447], and those considerations also apply to this document.

An experimental security protocol is defined in [PTPv2]. The PTP security extension and protocol provide group source authentication, message integrity, and replay attack protection for PTP messages.

9. References

9.1. Normative References


9.2. Informative References


Author’s Address

Ron Cohen
Resolute Networks Ltd.
Ligad Center P.O. Box 101
Modiin 71700
Israel

Email: ronc@resolutenetworks.com
Full Copyright Statement

Copyright (C) The IETF Trust (2007).

This document is subject to the rights, licenses and restrictions
contained in BCP 78, and except as set forth therein, the authors
retain all their rights.

This document and the information contained herein are provided on an
"AS IS" basis and THE CONTRIBUTOR, THE ORGANIZATION HE/SHE REPRESENTS
OR IS SPONSORED BY (IF ANY), THE INTERNET SOCIETY, THE IETF TRUST AND
THE INTERNET ENGINEERING TASK FORCE DISCLAIM ALL WARRANTIES, EXPRESS
OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF
THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED
WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Intellectual Property

The IETF takes no position regarding the validity or scope of any
Intellectual Property Rights or other rights that might be claimed to
pertain to the implementation or use of the technology described in
this document or the extent to which any license under such rights
might or might not be available; nor does it represent that it has
made any independent effort to identify any such rights. Information
on the procedures with respect to rights in RFC documents can be
found in BCP 78 and BCP 79.

Copies of IPR disclosures made to the IETF Secretariat and any
assurances of licenses to be made available, or the result of an
attempt made to obtain a general license or permission for the use of
such proprietary rights by implementers or users of this
specification can be obtained from the IETF on-line IPR repository at

The IETF invites any interested party to bring to its attention any
copyrights, patents or patent applications, or other proprietary
rights that may cover technology that may be required to implement
this standard. Please address the information to the IETF at
ietf-ipr@ietf.org.

Acknowledgment

Funding for the RFC Editor function is provided by the IETF
Administrative Support Activity (IASA).