TICTOC Working Group
Internet-Draft
Intended status: Standards Track
Expires: December 17, 2013

S. Davari
A. Oren
Broadcom Corp.
M. Bhatia
P. Roberts
Alcatel-Lucent
L. Montini
L. Martini
Cisco Systems
June 15, 2013

Transporting Timing messages over MPLS Networks
draft-ietf-tictoc-1588overmpls-05

Abstract

This document defines the method for transporting Timing messages such as PTP and NTP over an MPLS network. The method allows for the easy identification of these PDUs at the port level to allow for port level processing of these PDUs in both LERs and LSRs.

The basic idea is to transport Timing messages inside dedicated MPLS LSPs. These LSPs only carry Timing messages and possibly Control and Management packets, but they do not carry customer traffic.

Two methods for transporting Timing messages over MPLS are defined. The first method is to transport Timing messages directly over the dedicated MPLS LSP via UDP/IP encapsulation, which is suitable for MPLS networks. The second method is to transport Timing messages inside a PW via Ethernet encapsulation.

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at http://datatracker.ietf.org/drafts/current/.

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."

This Internet-Draft will expire on December 17, 2013.

Table of Contents

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 [RFC2119].

When used in lower case, these words convey their typical use in common language, and are not to be interpreted as described in RFC2119 [RFC2119].

1. Introduction

The objective of Precision Time Protocol (PTP) and Network Timing Protocol (NTP) are to synchronize independent clocks running on separate nodes of a distributed system.

[IEEE-1588] defines PTP messages for frequency, phase and time synchronization. The PTP messages include PTP PDUs over UDP/IP (Annex D and E of [IEEE-1588]) and PTP PDUs over Ethernet (Annex F of [IEEE-1588]). This document defines mapping and transport of the PTP messages defined in [IEEE-1588] over MPLS/MPLS-TP networks. PTP defines several clock types: ordinary clocks, boundary clocks, end-to-end transparent clocks, and peer-to-peer transparent clocks. Transparent clocks require intermediate nodes to update correction field inside PTP message that reflects the transit time in the node.

[RFC5905] defines NTP messages for clock and time synchronization. The PTP messages (PDUs) are transported over UDP/IP. This document defines mapping and transport of the NTP messages defined in [RFC5905] over MPLS networks.

One key attribute of all of these Timing messages is that the Time stamp processing should occur as close as possible to the actual transmission and reception at the physical port interface. This targets optimal time and/or frequency recovery by avoiding variable delay introduced by queues internal to the clocks.

To facilitate the fast and efficient recognition of Timing messages at the port level when the Timing messages are carried over MPLS LSPs, this document defines the specific encapsulations that should be used. In addition, it can be expected that there will exist LSR/LERs where only a subset of the physical ports will have the port-based Timing message processing capabilities. In order to ensure that the LSPs carrying Timing packets always enter and exit ports with this capability, routing extensions are defined to advertise this capability on a port basis and to allow for the establishment of LSPs that only transit such ports. While this path establishment restriction may be applied only at the LER Ingress and/or egress ports, it becomes more important when using transparent clock capable LSRs in the path.

Port based Timing message processing involves Timing message recognition. Once the Timing messages are recognized they can be modified based on the reception or transmission Time-stamp.

This document provides two methods for transporting Timing messages over MPLS. One is applicable to MPLS environment and the other one is applicable to MPLS/MPLS-TP environment

The solution involves transporting Timing messages over dedicated LSPs called Timing LSPs. These LSPs carry Timing messages and MAY carry Management and control messages, but not data plane client traffic. Timing LSPs can be established statically or via signaling. Extensions to control plane (OSPF, ISIS, etc.) is required to enable routers to distribute their Timing processing capabilities over MPLS to other routers. However such extensions are outside the scope of this document.

When signaling is used to setup the PTP LSP, Extensions to signaling protocols (e.g., RSVP-TE) are required for establishing PTP LSPs. However such extensions are outside the scope of this document.

While the techniques included herein allow for the establishment of paths optimized to include Time-stamping capable links, the performance of the Slave clocks is outside the scope of this document.

At the time of publishing this specification, Transparent Clocking (TC) is only defined for PTP. Therefore at this time any part of this specification that talks about Transparent Clocking applies only to PTP.

2. Terminology

1588: The timing and synchronization as defined by IEEE 1588.

NTP: The timing and synchronization protocol defined by IETF RFC-1305 and RFC-5905.

PTP: The timing and synchronization protocol used by 1588.

Master Clock: The source of 1588 timing to a set of slave clocks.

Master Port: A port on a ordinary or boundary clock that is in Master state. This is the source of timing toward slave ports.

Slave Clock: A receiver of 1588 timing from a master clock.

Slave Port: A port on a boundary clock or ordinary clock that is receiving timing from a master clock.

Ordinary Clock: A device with a single PTP port.

Transparent Clock. A device that measures the time taken for a PTP event message to transit the device and then updates the correctionField of the message with this transit time.

Boundary Clock: A device with more than one PTP port. Generally boundary clocks will have one port in slave state to receive timing and then other ports in master state to re-distribute the timing.

PTP LSP: An LSP dedicated to carry PTP messages

PTP PW: A PW within a PTP LSP that is dedicated to carry PTP messages.

CW: Pseudowire Control Word

LAG: Link Aggregation

ECMP: Equal Cost Multipath

CF: Correction Field, a field inside certain PTP messages (message type 0-3)that holds the accumulative transit time inside intermediate switches

Timing messages: Timing Protocol messages that are exchanged between routers in order to establish a synchronized clock.

3. Problem Statement

[IEEE-1588] has defined methods for transporting PTP messages over Ethernet and IP networks. [RFC5905] has defined the method of transporting NTP messages over IP networks. There is a need to transport Timing messages over MPLS networks while supporting the Transparent Clock (TC), Boundary Clock (BC) and Ordinary Clock (OC) functionality in the LER and LSRs in the MPLS network.

There are multiple ways of transporting Timing over MPLS. However, there is a requirement to limit the possible encapsulation options to simplify the Timing message identification and processing required at the port level.

When Timing-awareness is needed, Timing messages should not be transported over LSPs or PWs that are carrying customer traffic because LSRs perform Label switching based on the top label in the stack. To detect Timing messages inside such LSPs require special hardware to do deep packet inspection at line rate. Even if such hardware exists, the payload can't be deterministically identified by LSRs because the payload type is a context of the PW label, and the PW label and its context are only known to the Edge routers (PEs/ LERs); LSRs dont know what is a PWs payload (Ethernet, ATM, FR, CES, etc). Even if one restricts an LSP to only carry Ethernet PWs, the LSRs dont have the knowledge of whether PW Control Word (CW) is present or not and therefore can not deterministically identify the payload.

A generic method is defined in this document that does not require deep packet inspection at line rate, and can deterministically identify Timing messages. This method can be used to detect Timing Messages in both one-step and two-step clock implementations of ordinary, boundary and transparent clocks.

4. Timing over MPLS Architecture

Timing messages are exchange between Timing ports on ordinary and boundary clocks. Boundary clocks terminate the Timing messages and act as master for other boundary clocks or for slave clocks. End-to-End Transparent clocks do not terminate the Timing messages but they do modify the contents of the Timing messages as they transit across the transparent clock.

Master/Slave clocks (OCs), Boundary Clocks (BC) and Transparent Clock

(TC) could be implemented in either LERs or LSRs.

An example is shown in Figure 1, where the LERs act as Ordinary Clock (OC) and are the initiating/terminating point for Timing messages. The ingress LER encapsulates the Timing messages in Timing LSP and the Egress LER terminates the Timing LSP. The LSRs act as Transparent Clock (TC) and just update the Timing field in the Timing messages.

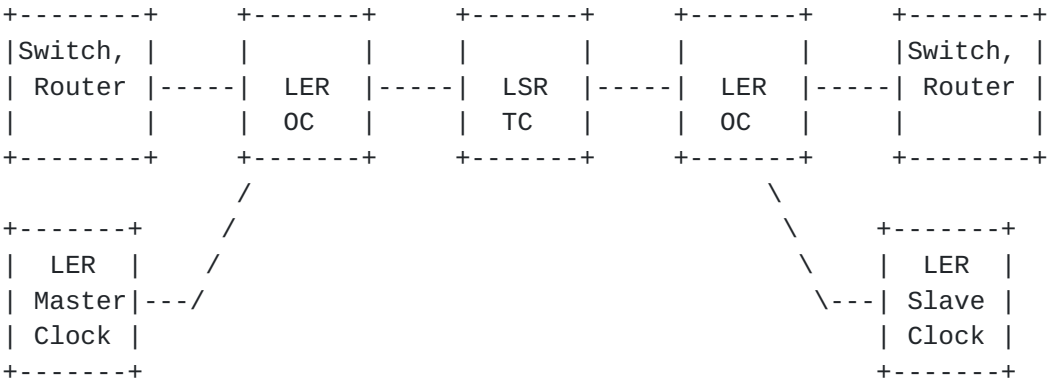

Figure (1) - Deployment example 1 of timing over MPLS network

Another example is shown in Figure2, where LERs terminate the Timing messages received from switch/routers that are outside of the MPLS network acting as OC or BC. In this example LERs regenerate the clock and initiate timing messages encapsulated in Timing LSP toward the MPLS network, while the LSRs act as Transparent Clock (TC) and just update the Timing field in the Timing messages, which are already encapsulated in Timing LSPs.

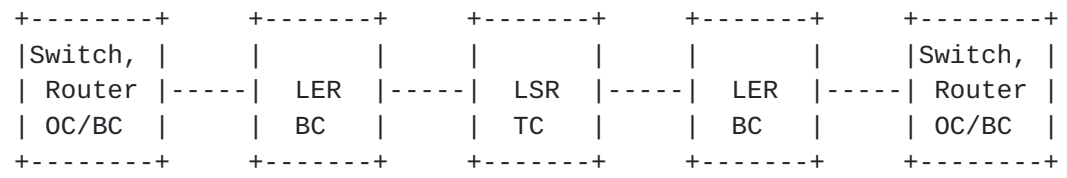

Figure (2) - Deployment example 2 of timing over MPLS network

Another example is shown in Figure 3, where LERs do not terminate the Timing messages received from switch/routers that are outside of the MPLS network acting as OC, TC or BC. The LERs act as TC and update the Timing field in the Timing messages as they transit the LER, while encapsulating them in timing LSP. The LSRs also act as Transparent Clock (TC) and just update the Timing field in the Timing messages which are already encapsulated in Timing LSPs.

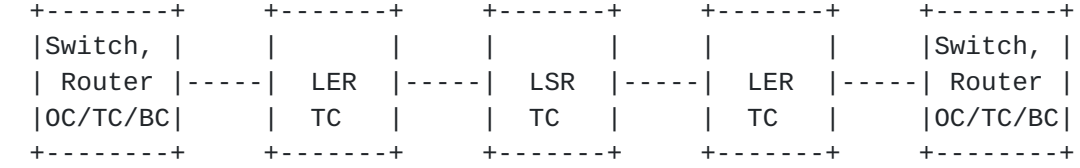

Figure (3) - Deployment example 3 of timing over MPLS network

Another example is shown in Figure 4, where LERs and LSRs support Boundary Clocks. A single-hop LSP is created between two adjacent LSRs engaged in BC operation. Other methods such as PTP transport over Ethernet MAY be used for transporting timing messages if the link between the two routers is Ethernet.

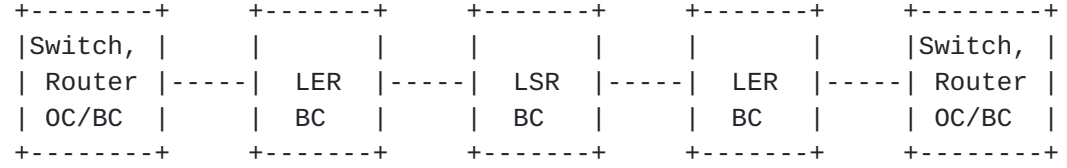

Figure (4) - Deployment example 3 of timing over MPLS network

An MPLS domain MAY serve multiple customers. In these cases the MPLS domain (maintained by a service provider) may provide timing services to multiple customers, each having their own Timing domain.

The Timing over MPLS architecture assumes full mesh of Timing LSPs between all LERs supporting this specification. It supports Point-to- point (VPWS) and Multipoint (VPLS) services. This means that a customer may purchase a Point-to-point Timing service between two customer sites or a Multipoint Timing service between more than

two customer sites.

The Timing over MPLS architecture supports P2P or P2MP Timing LSPs. This means that the Timing Multicast messages such as PTP Multicast event messages can be transported over P2MP Timing LSP or be replicated and transported over many P2P Timing LSPs.

Timing messages, that do not require Time stamping or Correction Field update MAY be transported over Timing LSPs to simplify hardware and software.

PTP Announce messages that determine the Timing LSP terminating point behavior such as BC/OC/TC SHOULD be transported over the Timing LSP to simplify hardware and software.

5. Dedicated LSPs for Timing messages

Many methods have been considered for identifying the Timing messages when they are encapsulated in MPLS such as using GAL/G-ACH or a new reserved label. These methods were not attractive since they either required deep packet inspection at line rate in the intermediate LSRs or they required use of a scarce new reserved label. Also one of the goals was to reuse existing OAM mechanisms.

The method defined in this document can be used by LER and LSRs to identify Timing messages in MPLS tunnels by just looking at the top label in the MPLS label stack, which only carry Timing messages as well as OAM, but not data plane client traffic.

Compliant implementations MUST use dedicated LSPs to carry Timing messages over MPLS. These LSPs are herein referred to as "Timing LSPs" and the labels associated with these LSPs as "Timing LSP labels". The Timing LSPs that runs between Ingress and Egress LERs MUST be co-routed. Alternatively, a single bidirectional co-routed LSP can be used.

Co-routing of the two directions is required to limit the difference in the delays in the Master clock to Slave clock direction compared to the Slave clock to Master clock direction. The Timing LSP MAY be MPLS/MPLS-TP LSP.

The Timing LSPs could be configured or signaled via RSVP-TE/GMPLS. New Extensions to RSVP-TE/GMPLS TLVs are required; however they are outside the scope of this document.

The Timing LSPs MAY carry essential MPLS/MPLS-TP OAM traffic such as BFD and LSP Ping but the LSP data plane client plane traffic MUST be Timing packets only.

6. Timing over LSP Encapsulation

This document defines two methods for carrying Timing messages over MPLS. The first method is carrying UDP/IP encapsulated Timing messages over Timing LSPs, and the second method, is carrying Ethernet encapsulated Timing messages over Ethernet PWs inside Timing LSPs.

6.1. Timing over UDP/IP over MPLS Encapsulation

The simplest method of transporting Timing messages over MPLS is to encapsulate Timing PDUs in UDP/IP and then encapsulate them in Timing LSP. This format is shown in Figure 4.

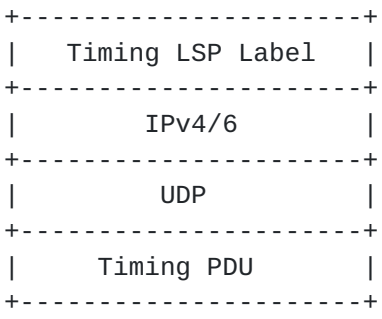

Figure (4) - Timing over UDP/IP over MPLS Encapsulation

This encapsulation is very simple and is useful when the network between Timing Master Clock and Slave Clock is MPLS network.

In order for an LER/LSR to process Timing messages, the Timing LSP Label must be at the top label of the label stack. The LER/LSR MUST know that the Timing LSP Label is used for carrying Timing messages. This can be accomplished via static configuration or via RSVP-TE signaling.

The UDP/IP encapsulation of PTP MUST follow Annex D and E of [IEEE-1588]. While the UDP/IP encapsulation of NTP MUST follow [RFC5905].

6.2. Timing over PW Encapsulation

Another method of transporting Timing over MPLS networks is by encapsulating Timing PDUs in PW which in turn is transported over Timing LSPs. In case of PTP, Ethernet PW encapsulation [RFC4448], shown in Fig 5(A) MUST be used and the Ethernet encapsulation of PTP MUST follow Annex F of [IEEE-1588].

The RAW mode or Tagged mode defined in [RFC4448] MAY be used and the Payload MUST have 0, 1, or 2 VLAN tags (S-VLAN and C-VLAN). The Timing over PW encapsulation MUST use the Control Word (CW) as specified in [RFC4448] to ensure proper detection of PTP messages inside the MPLS packets for Timing over LSP and Timing over PW encapsulation. The use of Sequence Number in the CW is optional.

Timing over PW encapsulation for NTP MUST use NTP over UDP/IP over PW (the IP PW discussed in [RFC4447]) shown in Fig 5(B).

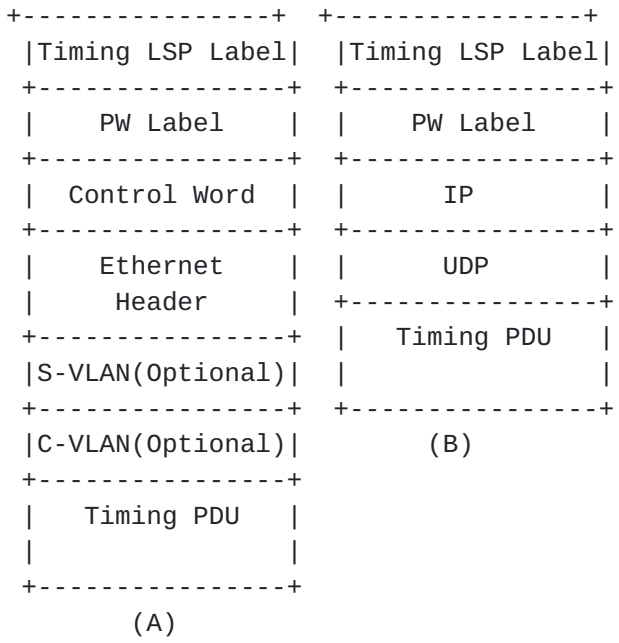

Figure (5) - Timing over PW Encapsulations

In order for an LSR to process PTP messages, the top label of the label stack (the Tunnel Label) MUST be a Timing label.

6.3. Other Timing Encapsulation methods

In future other timing encapsulation methods may be introduced, such as a new shim header after the Bottom of Stack to carry the Timing information. Such new encapsulations are outside the scope of this document.

7. Timing message Processing

Each Timing protocol such as PTP and NTP, define their set of Timing messages. For example PTP defines SYNC, DELAY_REQ, DELAY_RESP, FOLLOW_UP, etc messages.

Some of the Timing messages require time stamping or correction field update at port level and some dont. It is the job of the LER/LSR to parse the timing message and find out the type of the Timing message and decide whether and how to Time- stamp it (e.g., BC) or update correction field(e.g., TC).

For example the following PTP messages (called Event messages) require time-stamping or correction field update:

- SYNC
- DELAY_REQ (Delay Request)
- PDELAY_REQ (Peer Delay Request)
- PDELAY_RESP (Peer Delay Response)

SYNC and DELAY_REQ are exchanged between Master Clock and Slave Clock and MUST be transported over PTP LSPs. PDELAY_REQ and PDELAY_RESP are exchanged between adjacent PTP clocks (i.e. Master, Slave, Boundary, or Transparent) and SHOULD be transported over single hop PTP LSPs. If Two Step PTP clocks are present, then the FOLLOW_UP, and PDELAY_RESP_FOLLOW_UP messages MUST also be transported over the PTP LSPs.

For a given instance of 1588 protocol, SYNC and DELAY_REQ MUST be transported over two PTP LSPs that are in opposite directions. These PTP LSPs, which are in opposite directions MUST be congruent and co-routed. Alternatively, a single bidirectional co-routed LSP can be used.

Except as indicated above for the two-step PTP clocks, Non-Event PTP message types do not need to be processed by intermediate routers. These message types MAY be carried in PTP Tunnel LSPs.

8. Protection and Redundancy

In order to ensure continuous uninterrupted operation of slave clocks, usually as a general practice, slave clocks (or ports) track redundant master clocks.

It is the responsibility of the network operator to ensure that physically disjoint Timing LSPs are established between a slave clock (or port) and redundant master clocks (or ports).

When a slave clock (or port) listens to redundant master clocks or ports, any prolonged Timing LSP outage will trigger the slave clock or port to switch to a redundant master clock or port.

LSP/PW protection such as Linear protection Switching (1:1, 1+1), Ring protection switching or MPLS Fast Reroute (FRR) generally switch alternative path that usually cause a change in delay, which if undetected by slave clock can reduce accuracy of the slave clock.

Therefore protection switching MAY be used, as long as phase jumps upon switchover due to differences in path latency are detected and compensated for (such compensation not being required if BCs or peer-peer TCs are used throughout).

Note that any protection or reroute mechanism that adds additional MPLS label to the label stack, such as Facility Backup Fast Reroute, MUST ensure that the pushed label is also a Timing Label to ensure recognition of the MPLS frame as containing Timing messages, as it transits the backup path.

9. ECMP

To ensure the optimal operation of slave clocks and avoid error introduced by forward and reverse path delay asymmetry, the physical path for Timing messages from master clock to slave Clock and vice versa must be the same for all Event Timing messages listed in section 7.

Therefore the Timing LSPs MUST not be subject to ECMP (Equal Cost Multipath).

10. PHP

To ensure that the label on the top of the label stack is the Timing LSP Label, PHP MUST not be used.

11. Entropy

To ensure all Timing messages in a Timing LSP take the same path, Entropy Label MUST NOT be used for the Timing LSP[RFC6790] and Entropy Label MUST NOT be used for the PWs that are carried inside Timing LSP [RFC6391].

12. OAM, Control and Management

In order to monitor Timing LSPs and their encapsulated PWs, they MUST be able to carry OAM and management messages. These management messages MUST be differentiated from Timing messages via already defined IETF methods.

For example BFD [RFC5880], [RFC5884] and LSP-Ping [RFC4389] MAY run over PTP LSPs via UDP/IP encapsulation or via GAL/G-ACH. These Management protocols can easily be identified by the UDP Destination Port number or by GAL/G-ACH respectively.

Also BFD, LSP-Ping and other management messages MAY run over the PWs encapsulated in Timing LSP via one of the defined VCCVs (Type 1, 3 or 4) [RFC5085] (note that VCCV Type 2 using Router Alert Label is going to be deprecated by IETF). In this case G-ACH, PW label (TTL=1) or GAL-ACH are used to identify such management messages.

13. QoS Considerations

In network deployments where not every LSR/LER is Timing-aware, it is important to reduce the impact of the non-Timing-aware LSR/LERs on the timing recovery in the slave clock. The Timing messages are time critical and must be treated with the highest priority. Therefore Timing over MPLS messages must be treated with the highest priority in the routers. This can be achieved by proper setup of Timing LSPs.

It is recommended that the Timing LSPs are setup or configured properly to indicate EF-PHB [RFC3246]for the CoS and Green [RFC2697] for drop eligibility.

14. FCS and Checksum Recalculation

When time-stamp generation and timing packet adjustment is performed near the physical port hardware, the process MUST include recalculation of the Ethernet FCS. Also FCS retention for the payload Ethernet described in [RFC4720] MUST NOT be used.

For UDP/IP encapsulation mode of Timing over MPLS, the UDP checksum may be required as per UDP transport standards.

When UDP checksum is used, each Timing-aware LER/LSR must either incrementally update the UDP checksum after Time stamping or Correction Field update or verify the UDP checksum on reception from upstream and recalculate the checksum completely on transmission to downstream node after Time stamping or Correction Field update.

15. Behavior of LER/LSR

Timing-capable/aware LERs and LSRs are routers that have one or more interfaces that can perform Timing operations (OC/BC/TC) on Timing packets and are configured to do so. Timing-capable/aware LERs and LSRs can advertise their Timing-capability per-interface via control plane such as OSPF or IS-IS. The Timing-capable/aware LERs can then signals Timing LSPs via RSVP-TE signaling. Alternatively the Timing capability of LER and LSRs may be configured in a centralized controller and the Timing LSP may be setup using manual configuration or other methods such as SDN.

15.1. Behavior of Timing-capable/aware LER

When a Timing-capable/aware LER behaves as a Transparent clock and receives a Timing message from a Timing-capable/aware non-MPLS interface, the LER updates the Correction Field (CF) and encapsulates and forwards the timing message over previously established Timing LSP. Also when a Timing message is received from a Timing-capable/ aware MPLS interface, LER updates the Correction Filed (CF) and decapsulates the MPLS encapsulation and forwards the timing message to a non-MPLS interface.

When a Timing-capable/aware LER behaves as a Boundary clock and receives a Timing message from a Timing-capable/aware non MPLS interface, the LER Timestamps the Timing packet and sends it to the LERs Boundary clock processing module. Also when a Timing message is received from a Timing- capable/aware MPLS interface, the LER Timestamps the Timing packet and sends it to the LERs Boundary clock processing module.

When a Timing-capable/aware LER behaves as an Ordinary Clock toward the MPLS network, and receives a Timing message from a Timing-capable/aware MPLS interface, the LER Timestamps the Timing packet and sends it to the LERs Ordinary clock processing module.

15.2. Behavior of Timing-capable/aware LSR

When a Timing-capable/aware LSR behaves as a Transparent clock and receives a Timing message from a Timing-capable/aware MPLS interface, The LSR updates the Correction Filed (CF) and forwards the timing message over another MPLS interface.

When a Timing-capable/aware LSR behaves as a Boundary clock and receives a Timing message from a Timing-capable/aware MPLS interface. The LSR performs the functions of a Boundary Clock in terminating the received Timing message and re-generating a new timing message over another (or the same) MPLS interface.

15.3. Behavior of non-Timing-capable/aware LSR

It is most beneficial when all LSRs in the path of a Timing LSP be timing-Capable/aware LSRs. This would ensure the highest quality time and clock synchronization by Timing Slave Clocks. However, this specification does not mandate that all LSRs in path of a Timing LSP be Timing- capable/aware.

Non-Timing-capable/aware LSRs just switch the packets encapsulated in Timing LSPs and dont perform any Timing operation (TC or BC). However as explained in QoS section the Timing over MPLS packets MUST be still be treated with the highest priority based on their Traffic Class (TC) marking.

16. Other considerations

[IEEE-1588] defines an optional peer-to-peer Transparent clocking that requires peer delay measurement between two adjacent Timing-capable/ aware routers/switches. Peer delay measurement messages need to be time stamped and terminated by the Timing-capable/aware routers/ switches. This means that two adjacent LSRs may be engaged in a peer delay measurement.

For transporting such peer delay measurement messages a single-hop LSP SHOULD to be created between the two adjacent LSRs engaged in peer delay measurement to carry peer delay measurement messages. Other methods such as PTP transport over Ethernet MAY be used for transporting peer delay measurement messages if the link between the two routers is Ethernet.

In Peer-to-peer transparent clocking (P2P TC), a Timing-capable/ ware routers/switches MUST maintain a list of all the neighbors it needs to send a PDelay_Req to, where each neighbor corresponds to a timing LSP.

The use of Explicit Null Label (Label= 0 or 2) is acceptable as long as either the Explicit Null label is the bottom of stack label (applicable only to UDP/IP encapsulation) or the label below the Explicit Null label is a PTP label.

17. 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 [IEEE-1588].The PTP security extension and protocol provides group source authentication, message integrity, and replay attack protection for PTP messages.

When the MPLS network (provider network) serves multiple customers, it is important to maintain and process each customers clock and Timing messages separately from other customers to ensure there is no cross- customer effect. For example if an LER BC is synchronized to a specific grandmaster, belonging to customer A, then the LER MUST use that BC clock only for customer A to ensure that customer A cannot attack other customers by manipulating its time.

Timing messages MAY be encrypted or authenticated, provided that the LERs/LSRs that are Timing capable/aware can authenticate/ decrypt the timing messages.

18. Acknowledgements

The authors would like to thank Ron Cohen, Yaakov Stein, Tal Mizrahi, Stefano Ruffini, Peter Meyer, and other members of IETF for reviewing and providing feedback on this draft.

19. IANA Considerations

There are no IANA requirements in this specification.

20. References

20.1. Normative References

[IEEE-1588] IEEE 1588-2008, "IEEE Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems".

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

[RFC3985] Bryant, S. and P. Pate, "Pseudo Wire Emulation Edge-to-Edge (PWE3) Architecture", RFC 3985, March 2005.

[RFC4389] Thaler, D., Talwar, M., and C. Patel, "Neighbor Discovery Proxies (ND Proxy)", RFC 4389, April 2006.

[RFC4447] Martini, L., Rosen, E., El-Aawar, N., Smith, T., and G. Heron, "Pseudowire Setup and Maintenance Using the Label Distribution Protocol (LDP)", RFC 4447, April 2006.

[RFC4448] Martini, L., Rosen, E., El-Aawar, N., and G. Heron, "Encapsulation Methods for Transport of Ethernet over MPLS Networks", RFC 4448, April 2006.

[RFC4720] Malis, A., Allan, D., and N. Del Regno, "Pseudowire Emulation Edge-to-Edge (PWE3) Frame Check Sequence Retention", RFC 4720, November 2006.

[RFC5085] Nadeau, T. and C. Pignataro, "Pseudowire Virtual Circuit Connectivity Verification (VCCV): A Control Channel for Pseudowires", RFC 5085, December 2007.

[RFC5880] Katz, D. and D. Ward, "Bidirectional Forwarding Detection (BFD)", RFC 5880, June 2010.

[RFC5884] Aggarwal, R., Kompella, K., Nadeau, T., and G. Swallow, "Bidirectional Forwarding Detection (BFD) for MPLS Label Switched Paths (LSPs)", RFC 5884, June 2010.

20.2. Informative References

[I-D.ietf-pwe3-fat-pw] Bryant, S., Filsfils, C., Drafz, U., Kompella, V., Regan, J., and S. Amante, "Flow Aware Transport of Pseudowires over an MPLS Packet Switched Network", draft-ietf-pwe3-fat-pw-07 (work in progress), July 2011.

[ISO] ISO/IEC 10589:1992, "Intermediate system to Intermediate system routeing information exchange protocol for use in conjunction with the Protocol for providing the Connectionless-mode Network Service (ISO 8473)".

[RFC1195] Callon, R., "Use of OSI IS-IS for routing in TCP/IP and dual environments", RFC 1195, December 1990.

[RFC2328] Moy, J., "OSPF Version 2", STD 54, RFC 2328, April 1998.

[RFC2697] Heinanen, J. and R. Guerin, "A Single Rate Three Color Marker", RFC 2697, September 1999.

[RFC3246] Davie, B., Charny, A., Bennet, J., Benson, K., Le Boudec, J., Courtney, W., Davari, S., Firoiu, V., and D. Stiliadis, "An Expedited Forwarding PHB (Per-Hop Behavior)", RFC 3246, March 2002.

[RFC3630] Katz, D., Kompella, K., and D. Yeung, "Traffic Engineering (TE) Extensions to OSPF Version 2", RFC 3630, September 2003.

[RFC3784] Smit, H. and T. Li, "Intermediate System to Intermediate System (IS-IS) Extensions for Traffic Engineering (TE)", RFC 3784, June 2004.

[RFC4970] Lindem, A., Shen, N., Vasseur, JP., Aggarwal, R., and S. Shaffer, "Extensions to OSPF for Advertising Optional Router Capabilities", RFC 4970, July 2007.

[RFC4971] Vasseur, JP., Shen, N., and R. Aggarwal, "Intermediate System to Intermediate System (IS-IS) Extensions for Advertising Router Information", RFC 4971, July 2007.

[RFC5120] Przygienda, T., Shen, N., and N. Sheth, "M-ISIS: Multi Topology (MT) Routing in Intermediate System to Intermediate Systems (IS-ISs)", RFC 5120, February 2008.

[RFC5305] Li, T. and H. Smit, "IS-IS Extensions for Traffic Engineering", RFC 5305, October 2008.

[RFC5329] Ishiguro, K., Manral, V., Davey, A., and A. Lindem, "Traffic Engineering Extensions to OSPF Version 3", RFC 5329, September 2008.

[RFC5340] Coltun, R., Ferguson, D., Moy, J., and A. Lindem, "OSPF for IPv6", RFC 5340, July 2008.

[RFC5905] Mills, D., Martin, J., Burbank, J., and W. Kasch, "Network Time Protocol Version 4: Protocol and Algorithms Specification", RFC 5905, June 2010.

[RFC6391] Bryant, S., Filsfils, C., Drafz, U., Kompella, V., Regan, J., and S. Amante, "Flow-Aware Transport of Pseudowires over an MPLS Packet Switched Network", RFC 6391, November 2011.

[RFC6790] Kompella, K., Drake, J., Amante, S., Henderickx, W., and L. Yong, "The Use of Entropy Labels in MPLS Forwarding", RFC 6790, November 2012.

1. Routing extensions for Timing-aware Routers

MPLS-TE routing relies on extensions to OSPF [RFC2328] [RFC5340] and IS-IS [ISO] [RFC1195] in order to advertise Traffic Engineering (TE) link information used for constraint-based routing.

Indeed, it is useful to advertise data plane TE router link capabilities, such as the capability for a router to be Timing-aware. This capability MUST then be taken into account during path computation to prefer or even require links that advertise themselves as Timing-aware. In this way the path can ensure the entry and exit points into the LERs and, if desired, the links into the LSRs are able to perform port based time-stamping thus minimizing their impact on the performance of the slave clock.

extensions are required to OSPF and IS-IS in order to advertise Timing-aware capabilities of a link. Such extensions are outside the scope of this document; however such extension SHOULD be able to signal the following information per Router Link:

- Capable of processing PTP, NTP or other Timing flows
- Capable of performing Transparent Clock operation
- Capable of performing Boundary Clock operation

2. Signaling Extensions for Creating Timing LSPs

RSVP-TE signaling MAY be used to setup the timing LSPs. When RSVP-TE is used to setup Timing LSPs, some information that indicates that the LSP is carrying Timing flows MUST be included in the new Extensions to RSVP-TE:

The following information MAY also be included in the new Extensions to RSVP-TE:

- Offset from Bottom of Stack (BoS) to the start of the Time-stamp field
- Number of VLANs in case of PW encapsulation
- Timestamp field Type
  - Correction Field, Timestamp
- Timestamp Field format
  - 64-bit PTPv1, 80-bit PTPv2, 32-bit NTP, 64-bit NTP, 128-bit NTP, etc.

Note that in case the above optional information is signaled with RSVP-TE for a Timing LSP, all the Timing packets carried in that LSP must have the same signaled characteristics. For example if Timestamp format is signaled as 64-bit PTPv1, then all Timing packets must use 64-bit PTPv1 time-stamp.

Authors' Addresses

Shahram Davari
Broadcom Corp.
San Jose, CA 95134
USA

Email: davari@broadcom.com

Amit Oren
Broadcom Corp.
San Jose, CA 95134
USA

Email: amito@broadcom.com

Manav Bhatia
Alcatel-Lucent
Bangalore,
India

Email: manav.bhatia@alcatel-lucent.com

Peter Roberts
Alcatel-Lucent
Kanata,
Canada

Email: peter.roberts@alcatel-lucent.com

Laurent Montini
Cisco Systems
San Jose CA
USA

Email: lmontini@cisco.com

Luca
Cisco Systems
San Jose CA
USA

Email: lmartini@cisco.com