

MPLS Working Group
Internet-Draft
Intended status: Standards Track
Expires: October 30, 2016

G. Mirsky
S. Ruffini
E. Gray
Ericsson
J. Drake
Juniper Networks
S. Bryant
Cisco Systems
A. Vainshtein
ECI Telecom
April 28, 2016

**Residence Time Measurement in MPLS network
draft-ietf-mpls-residence-time-08**

Abstract

This document specifies G-ACh based Residence Time Measurement and how it can be used by time synchronization protocols being transported over MPLS domain.

Residence time is the variable part of propagation delay of timing and synchronization messages and knowing what this delay is for each message allows for a more accurate determination of the delay to be taken into account in applying the value included in a PTP event message.

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 October 30, 2016.

Copyright Notice

Copyright (c) 2016 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to [BCP 78](#) and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1.	Introduction	3
1.1.	Conventions used in this document	3
1.1.1.	Terminology	3
1.1.2.	Requirements Language	4
2.	Residence Time Measurement	4
3.	G-ACh for Residence Time Measurement	5
3.1.	PTP Packet Sub-TLV	6
4.	Control Plane Theory of Operation	7
4.1.	RTM Capability	7
4.2.	RTM Capability Sub-TLV	8
4.3.	RTM Capability Advertisement in OSPFv2	9
4.4.	RTM Capability Advertisement in OSPFv3	9
4.5.	RTM Capability Advertisement in IS-IS	9
4.6.	RSVP-TE Control Plane Operation to Support RTM	10
4.7.	RTM_SET TLV	11
4.7.1.	RTM_SET Sub-TLVs	13
5.	Data Plane Theory of Operation	16
6.	Applicable PTP Scenarios	16
7.	One-step Clock and Two-step Clock Modes	17
8.	IANA Considerations	19
8.1.	New RTM G-ACh	19
8.2.	New RTM TLV Registry	19
8.3.	New RTM Sub-TLV Registry	20
8.4.	RTM Capability sub-TLV in OSPFv2	20
8.5.	RTM Capability sub-TLV in OSPFv3	21
8.6.	IS-IS RTM Application ID	21
8.7.	RTM_SET Sub-object RSVP Type and sub-TLVs	21
8.8.	RTM_SET Attribute Flag	22
8.9.	New Error Codes	23
9.	Security Considerations	23
10.	Acknowledgements	24

11.	References	24
11.1.	Normative References	24
11.2.	Informative References	25
	Authors' Addresses	26

[1.](#) Introduction

Time synchronization protocols, e.g., Network Time Protocol version 4 (NTPv4) [[RFC5905](#)] and Precision Time Protocol (PTP) Version 2 [[IEEE.1588.2014](#)] define timing messages that can be used to synchronize clocks across a network domain. Measurement of the cumulative time one of these timing messages spends transiting the nodes on the path from ingress node to egress node is termed Residence Time and it is used to improve the accuracy of clock synchronization. (I.e., it is the sum of the difference between the time of receipt at an ingress interface and the time of transmission from an egress interface for each node along the path from ingress node to egress node.) This document defines a new Generalized Associated Channel (G-ACh) value and an associated residence time measurement (RTM) packet that can be used in a Multi-Protocol Label Switching (MPLS) network to measure residence time over a Label Switched Path (LSP).

Although it is possible to use RTM over an LSP instantiated using LDP, that is outside the scope of this document. Rather, this document describes RTM over an LSP signaled using RSVP-TE [[RFC3209](#)] because the LSP's path can be either explicitly specified or determined during signaling.

Comparison with alternative proposed solutions such as [[I-D.ietf-tictoc-1588overmpls](#)] is outside the scope of this document.

[1.1.](#) Conventions used in this document

[1.1.1.](#) Terminology

MPLS: Multi-Protocol Label Switching

ACH: Associated Channel

TTL: Time-to-Live

G-ACh: Generic Associated Channel

GAL: Generic Associated Channel Label

NTP: Network Time Protocol

ppm: parts per million

PTP: Precision Time Protocol

LSP: Label Switched Path

OAM: Operations, Administration, and Maintenance

RR0: Record Route Object

RTM: Residence Time Measurement

IGP: Internal Gateway Protocol

1.1.2. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [\[RFC2119\]](#).

2. Residence Time Measurement

Packet Loss and Delay Measurement for MPLS Networks [\[RFC6374\]](#) can be used to measure one-way or two-way end-to-end propagation delay over LSP or PW. But these measurements are insufficient for use in some applications, for example, time synchronization across a network as defined in the Precision Time Protocol (PTP). In PTPv2 [\[IEEE.1588.2014\]](#) residence times is accumulated in the correctionField of the PTP event message, as defined in [\[IEEE.1588.2014\]](#), or in the associated follow-up message (or Delay_Resp message associated with the Delay_Req message) in case of two-step clocks (see the detailed discussion in [Section 7](#)).

IEEE 1588 uses this residence time to correct the transit time from ingress node to egress node, effectively making the transit nodes transparent.

This document proposes a mechanism that can be used as one of types of on-path support for a clock synchronization protocol or to perform one-way measurement of residence time. The proposed mechanism accumulates residence time from all nodes that support this extension along the path of a particular LSP in Scratch Pad field of an RTM packet Figure 1. This value can then be used by the egress node to update, for example, the correctionField of the PTP event packet carried within the RTM packet prior to performing its PTP processing.

3. G-ACh for Residence Time Measurement

[RFC 5586](#) [[RFC5586](#)] and [RFC 6423](#) [[RFC6423](#)] define the G-ACh to extend the applicability of the PW Associated Channel (ACH) [[RFC5085](#)] to LSPs. G-ACh provides a mechanism to transport OAM and other control messages over an LSP. Processing of these messages by select transit nodes is controlled by the use of the Time-to-Live (TTL) value in the MPLS header of these messages.

The packet format for Residence Time Measurement (RTM) is presented in Figure 1

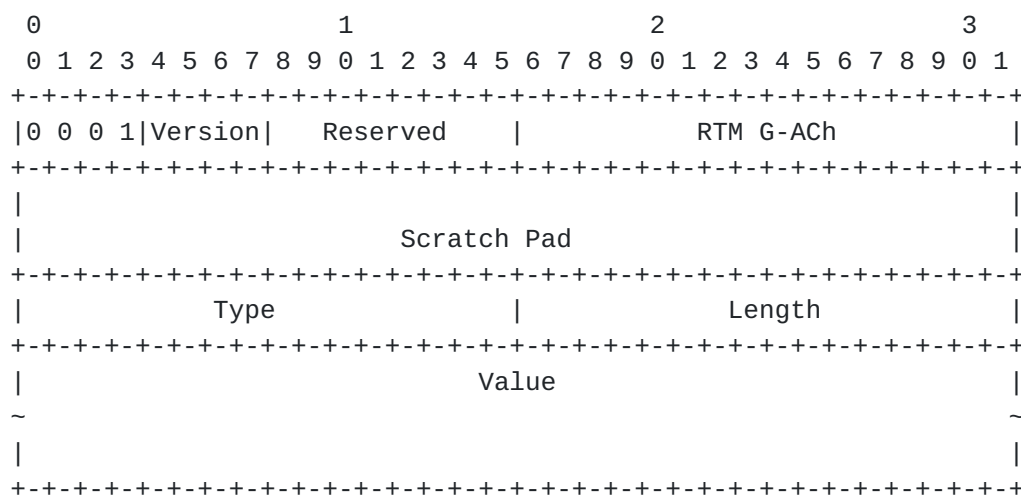


Figure 1: RTM G-ACh packet format for Residence Time Measurement

- o First four octets are defined as G-ACh Header in [[RFC5586](#)]
- o The Version field is set to 0, as defined in [RFC 4385](#) [[RFC4385](#)].
- o The Reserved field MUST be set to 0 on transmit and ignored on receipt.
- o The RTM G-ACh field, value (TBA1) to be allocated by IANA, identifies the packet as such.
- o The Scratch Pad field is 8 octets in length. It is used to accumulate the residence time spent in each RTM capable node transited by the packet on its path from ingress node to egress node. The first RTM-capable node MUST initialize the Scratch Pad field with its residence time measurement. Its format is IEEE double precision and its units are nanoseconds. Note that depending on whether the timing procedure is one-step or two-step operation ([Section 7](#)), the residence time is either for the timing packet carried in the Value field of this RTM packet or for an

associated timing packet carried in the Value field of another RTM packet.

- o The Type field identifies the type and encapsulation of a timing packet carried in the Value field, e.g., NTP [[RFC5905](#)] or PTP [[IEEE.1588.2014](#)]. IANA will be asked to create a sub-registry in Generic Associated Channel (G-ACh) Parameters Registry called "MPLS RTM TLV Registry".
- o The Length field contains the length, in octets , of the of the timing packet carried in the Value field.
- o The optional Value field MAY carry a packet of the time synchronization protocol identified by Type field. It is important to note that the packet may be authenticated or encrypted and carried over LSP edge to edge unchanged while the residence time is accumulated in the Scratch Pad field.
- o The TLV MUST be included in the RTM message, even if the length of the Value field is zero.

[3.1.](#) PTP Packet Sub-TLV

Figure 2 presents format of a PTP sub-TLV that MUST be included in the Value field of an RTM packet preceding the carried timing packet when the timing packet is PTP.



Figure 2: PTP Sub-TLV format

where Flags field has format

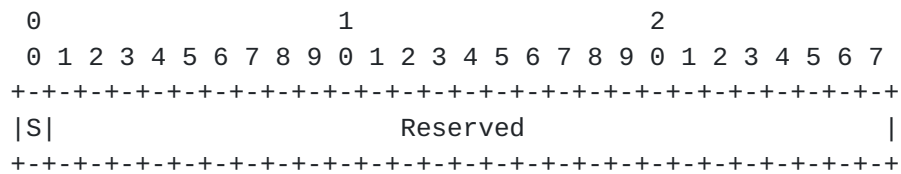


Figure 3: Flags field format of PTP Packet Sub-TLV

- o The Type field identifies PTP sub-TLV defined in the Table 19 Values of messageType field in [[IEEE.1588.2014](#)].
- o The Length field of the PTP sub-TLV contains the number of octets of the Value field and MUST be 20.
- o The Flags field currently defines one bit, the S-bit, that defines whether the current message has been processed by a 2-step node, where the flag is cleared if the message has been handled exclusively by 1-step nodes and there is no follow-up message, and set if there has been at least one 2-step node and a follow-up message is forthcoming.
- o The PTPTYPE indicates the type of PTP packet carried in the TLV. PTPTYPE is the messageType field of the PTPv2 packet whose values are defined in the Table 19 [[IEEE.1588.2014](#)].
- o The 10 octets long Port ID field contains the identity of the source port.
- o The Sequence ID is the sequence ID of the PTP message carried in the Value field of the message.

4. Control Plane Theory of Operation

The operation of RTM depends upon TTL expiry to deliver an RTM packet from one RTM capable interface to the next along the path from ingress node to egress node. This means that a node with RTM capable interfaces MUST be able to compute a TTL which will cause the expiry of an RTM packet at the next node with RTM capable interfaces.

4.1. RTM Capability

Note that the RTM capability of a node is with respect to the pair of interfaces that will be used to forward an RTM packet. In general, the ingress interface of this pair must be able to capture the arrival time of the packet and encode it in some way such that this information will be available to the egress interface.

The supported modes (1-step verses 2-step) of any pair of interfaces is then determined by the capability of the egress interface. For both modes, the egress interface implementation **MUST** be able to determine the precise departure time of the same packet and determine from this, and the arrival time information from the corresponding ingress interface, the difference representing the residence time for the packet.

An interface with the ability to do this and update the associated Scratch Pad in real-time (i.e. while the packet is being forwarded) is said to be 1-step capable.

Hence while both ingress and egress interfaces are required to support RTM for the pair to be RTM-capable, it is the egress interface that determines whether or not the node is 1-step or 2-step capable with respect to the interface-pair.

The RTM capability used in the sub-TLV shown in Figure 4 is thus associated with the egress port of the node making the advertisement, while the ability of any pair of interfaces that includes this egress interface to support any mode of RTM depends on the ability of that interface to record packet arrival time in some way that can be conveyed to and used by that egress interface.

When a node uses an IGP to carry the RTM capability sub-TLV, the sub-TLV **MUST** reflect the RTM capability (1-step or 2-step) associated with egress interfaces.

[4.2.](#) RTM Capability Sub-TLV

The format for the RTM Capabilities sub-TLV is presented in Figure 4

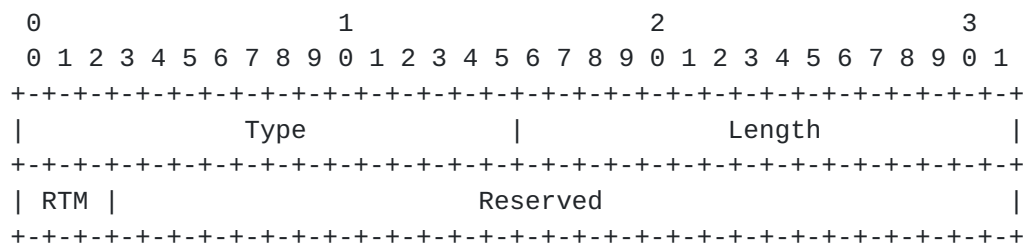


Figure 4: RTM Capability sub-TLV

- o Type values TBA2 and TBA3 will be assigned by IANA from appropriate registries for OSPFv2 and OSPFv3 respectively.
- o Length **MUST** be set to 4.

- o RTM (capability) - is a three-bit long bit-map field with values defined as follows:
 - * 0b001 - one-step RTM supported;
 - * 0b010 - two-step RTM supported;
 - * 0b100 - reserved.
- o Reserved field must be set to all zeroes on transmit and ignored on receipt.

[RFC4202] explains that the Interface Switching Capability Descriptor describes switching capability of an interface. For bi-directional links, the switching capabilities of an interface are defined to be the same in either direction. I.e., for data entering the node through that interface and for data leaving the node through that interface. That principle SHOULD be applied when a node advertises RTM Capability.

A node that supports RTM MUST be able to act in two-step mode and MAY also support one-step RTM mode. Detailed discussion of one-step and two-step RTM modes in [Section 7](#).

[4.3.](#) RTM Capability Advertisement in OSPFv2

The capability to support RTM on a particular link (interface) is advertised in the OSPFv2 Extended Link Opaque LSA described in [Section 3 \[RFC7684\]](#) via the RTM Capability sub-TLV.

Its Type value will be assigned by IANA from the OSPF Extended Link TLV Sub-TLVs registry that will be created per [\[RFC7684\]](#) request.

[4.4.](#) RTM Capability Advertisement in OSPFv3

The capability to support RTM on a particular link (interface) is advertised in the OSPFv3 be Intra-Area-Prefix TLV, IPv6 Link-Local Address TLV, or the IPv4 Link-Local Address TLV described in [\[I-D.ietf-ospf-ospfv3-lsa-extend\]](#) via the RTM Capability sub-TLV.

[4.5.](#) RTM Capability Advertisement in IS-IS

The capability to support RTM on a particular link (interface) is advertised in the GENINFO TLV described in [\[RFC6823\]](#) via the RTM Capability sub-TLV.

With respect to the Flags field of the GENINFO TLV:

- o The S bit MUST be cleared to prevent the RTM Capability sub-TLV from leaking between levels.
- o The D bit of the Flags field MUST be cleared as required by [\[RFC6823\]](#).
- o The I bit and the V bit MUST be set accordingly depending on whether RTM capability being advertised is for an IPv4 or an IPv6 interface.

Application ID (TBA4) will be assigned from the Application Identifiers for TLV 251 IANA registry. The RTM Capability sub-TLV MUST be included in GENINFO TLV in Application Specific Information.

4.6. RSVP-TE Control Plane Operation to Support RTM

Throughout this document we refer to a node as RTM capable node when at least one of its interfaces is RTM capable. Figure 5 provides an example of roles a node may have with respect to RTM capability:



Figure 5: RTM capable roles

- o A is a Boundary Clock with its egress port in Master state. Node A transmits IP encapsulated timing packets whose destination IP address is G.
- o B is the ingress LER for the MPLS LSP and is the first RTM capable node. It creates RTM packets and in each it places a timing packet, possibly encrypted, in the Value field and initializes the Scratch Pad field with its residence time measurement
- o C is a transit node that is not RTM capable. It forwards RTM packets without modification.
- o D is RTM capable transit node. It updates the Scratch Pad field of the RTM packet without updating of the timing packet.
- o E is a transit node that is not RTM capable. It forwards RTM packets without modification.
- o F is the egress LER and the last RTM capable node. It processes the timing packet carried in the Value field using the value in the Scratch Pad field. It updates the Correction field of the PTP

message with the value in the Scratch Pad field of the RTM ACH, and removes the RTM ACH encapsulation.

- o G is a Boundary Clock with its ingress port in Slave state. Node G receives PTP messages.

An ingress node that is configured to perform RTM along a path through an MPLS network to an egress node verifies that the selected egress node has an interface that supports RTM via the egress node's advertisement of the RTM Capability sub-TLV. In the Path message that the ingress node uses to instantiate the LSP to that egress node it places LSP_ATTRIBUTES Object [[RFC5420](#)] with RTM_SET Attribute Flag set [Section 8.8](#) which indicates to the egress node that RTM is requested for this LSP. RTM_SET Attribute Flag SHOULD NOT be set in the LSP_REQUIRED_ATTRIBUTES object [[RFC5420](#)] , unless it is known that all nodes support RTM, because a node that does not recognize RTM_SET Attribute Flag would reject the Path message.

If egress node receives Path message with RTM_SET Attribute Flag in LSP_ATTRIBUTES object, it MUST include initialized RRO [[RFC3209](#)] and LSP_ATTRIBUTES object where RTM_SET Attribute Flag is set and RTM_SET TLV [Section 4.7](#) is initialized. When Resv message received by ingress node the RTM_SET TLV will contain an ordered list, from egress node to ingress node, of the RTM capable node along the LSP's path.

After the ingress node receives the Resv, it MAY begin sending RTM packets on the LSP's path. Each RTM packet has its Scratch Pad field initialized and its TTL set to expire on the closest downstream RTM capable node.

It should be noted that RTM can also be used for LSPs instantiated using [[RFC3209](#)] in an environment in which all interfaces in an IGP support RTM. In this case the RTM_SET TLV and LSP_ATTRIBUTES Object MAY be omitted.

[4.7.](#) RTM_SET TLV

RTM capable interfaces can be recorded via RTM_SET TLV. The RTM_SET sub-object format is of generic Type, Length, Value (TLV), presented in Figure 6 .

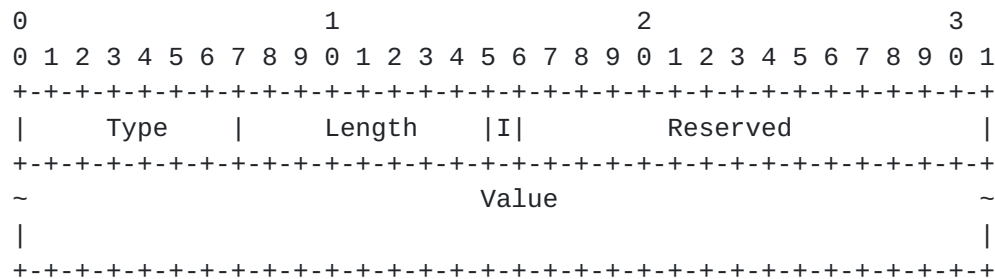


Figure 6: RTM_SET TLV format

Type value (TBA5) will be assigned by IANA from its Attributes TLV Space sub-registry.

The Length contains the total length of the sub-object in bytes, including the Type and Length fields.

The I bit flag indicates whether the downstream RTM capable node along the LSP is present in the RRO.

Reserved field must be zeroed on initiation and ignored on receipt.

The content of an RTM_SET TLV is a series of variable-length sub-TLVs. Only a single RTM_SET can be present in the LSP_ATTRIBUTES object. The sub-TLVs are defined in [Section 4.7.1](#) below.

The following processing procedures apply to every RTM capable node along the LSP that in this paragraph is referred as node for sake of brevity. Each node MUST examine Resv message whether RTM_SET Attribute Flag in the LSP_ATTRIBUTES object is set. If the RTM_SET flag set, the node MUST inspect the LSP_ATTRIBUTES object for presence of RTM_SET TLV. If more than one found, then the LSP setup MUST fail with generation of the ResvErr message with Error Code Duplicate TLV [Section 8.9](#) and Error Value that contains Type value in its 8 least significant bits. If no RTM_SET TLV has been found, then the LSP setup MUST fail with generation of the ResvErr message with Error Code RTM_SET TLV Absent [Section 8.9](#). If one RTM_SET TLV has been found the node will use the ID of the first node in the RTM_SET in conjunction with the RRO to compute the hop count to its downstream node with reachable RTM capable interface. If the node cannot find matching ID in RRO, then it MUST try to use ID of the next node in the RTM_SET until it finds the match or reaches the end of RTM_SET TLV. If match have been found, then the calculated value is used by the node as TTL value in outgoing label to reach the next RTM capable node on the LSP. Otherwise, the TTL value MUST be set to 255. The node MUST add RTM_SET sub-TLV with the same address it used in RRO sub-object at the beginning of the RTM_SET TLV in associated outgoing Resv message before forwarding it upstream. If the

calculated TTL value been set to 255, as described above, then the I flag in node RTM_SET TLV MUST be set to 1 before Resv message forwarded upstream. Otherwise, the I flag MUST be cleared (0).

The ingress node MAY inspect the I bit flag received in each RTM_SET TLV contained in the LSP_ATTRIBUTES object of a received Resv message. Presence of the RTM_SET TLV with I bit field set to 1 indicates that some RTM nodes along the LSP could be included in the calculation of the residence time. An ingress node MAY choose to resignal the LSP to include all RTM nodes or simply notify the user via a management interface.

There are scenarios when some information is removed from an RRO due to policy processing (e.g., as may happen between providers) or RRO is limited due to size constraints. Such changes affect the core assumption of the method to control processing of RTM packets. RTM SHOULD NOT be used if it is not guaranteed that RRO contains complete information.

4.7.1. RTM_SET Sub-TLVs

The RTM Set sub-object contains an ordered list, from egress node to ingress node, of the RTM capable nodes along the LSP's path.

The contents of a RTM_SET sub-object are a series of variable-length sub-TLVs. Each sub-TLV has its own Length field. The Length contains the total length of the sub-TLV in bytes, including the Type and Length fields. The Length MUST always be a multiple of 4, and at least 8 (smallest IPv4 sub-object).

Sub-TLVs are organized as a last-in-first-out stack. The first -out sub-TLV relative to the beginning of RTM_SET TLV is considered the top. The last-out sub-TLV is considered the bottom. When a new sub-TLV is added, it is always added to the top. Only a single RTM_SET sub-TLV with the given Value field MUST be present in the RTM_SET TLV. If more than one sub-TLV is found the LSP setup MUST fail with the generation of a ResvErr message with the Error Code "Duplicate sub-TLV" [Section 8.9](#) and Error Value contains 16-bit value composed of (Type of TLV, Type of sub-TLV).

Three kinds of sub-TLVs for RTM_SET are currently defined.

4.7.1.1. IPv4 Sub-TLV

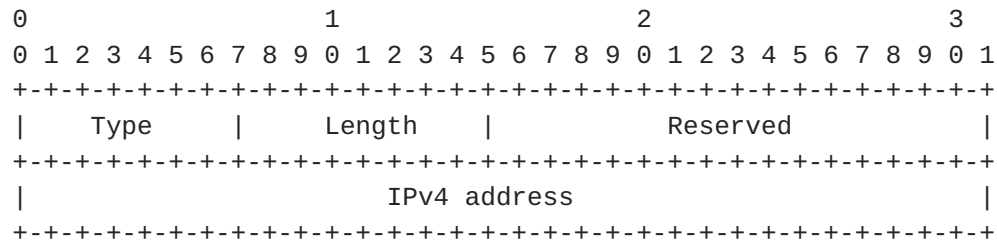


Figure 7: IPv4 sub-TLV format

Type

0x01 IPv4 address

Length

The Length contains the total length of the sub-TLV in bytes, including the Type and Length fields. The Length is always 8.

IPv4 address

A 32-bit unicast host address.

Reserved

Zeroed on initiation and ignored on receipt.

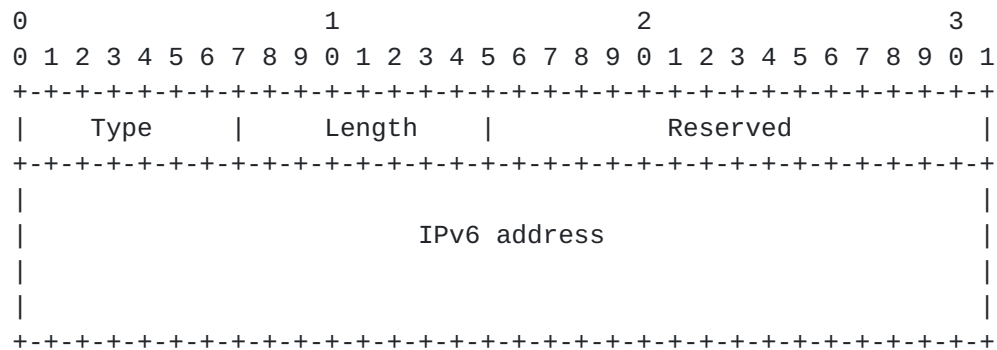
4.7.1.2. IPv6 Sub-TLV

Figure 8: IPv6 sub-TLV format

Type

0x02 IPv6 address

Length

The Length contains the total length of the sub-TLV in bytes, including the Type and Length fields. The Length is always 20.

IPv6 address

A 128-bit unicast host address.

Reserved

Zeroed on initiation and ignored on receipt.

4.7.1.3. Unnumbered Interface Sub-TLV

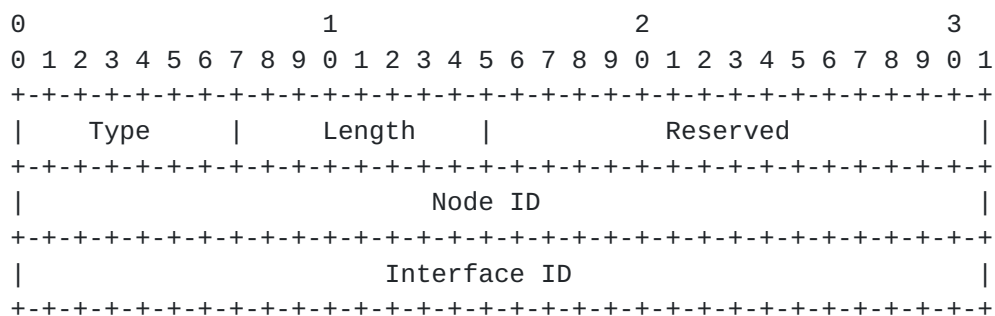


Figure 9: IPv4 sub-TLV format

Type

0x03 Unnumbered interface

Length

The Length contains the total length of the sub-TLV in bytes, including the Type and Length fields. The Length is always 12.

Node ID

The Node ID interpreted as Router ID as discussed in the [Section 2 \[RFC3477\]](#).

Interface ID

The identifier assigned to the link by the node specified by the Node ID.

Reserved

Zeroed on initiation and ignored on receipt.

5. Data Plane Theory of Operation

After instantiating an LSP for a path using RSVP-TE [[RFC3209](#)] as described in [Section 4.6](#) or as described in the second paragraph of [Section 4](#) and in [Section 4.6](#), ingress node MAY begin sending RTM packets to the first downstream RTM capable node on that path. Each RTM packet has its Scratch Pad field initialized and its TTL set to expire on the next downstream RTM-capable node. Each RTM-capable node on the explicit path receives an RTM packet and records the time at which it receives that packet at its ingress interface as well as the time at which it transmits that packet from its egress interface; this should be done as close to the physical layer as possible to ensure precise accuracy in time determination. The RTM-capable node determines the difference between those two times; for 1-step operation, this difference is determined just prior to or while sending the packet, and the RTM-capable egress interface adds it to the value in the Scratch Pad field of the message in progress. Note, for the purpose of calculating a residence time, a common free running clock synchronizing all the involved interfaces may be sufficient, as, for example, 4.6 ppm accuracy leads to 4.6 nanosecond error for residence time on the order of 1 millisecond.

For 2-step operation, the difference between packet arrival time (at an ingress interface) and subsequent departure time (from an egress interface) is determined at some later time prior to sending a subsequent follow-up message, so that this value can be used to update the correctionField in the follow-up message.

See [Section 7](#) for further details on the difference between 1-step and 2-step operation.

The last RTM-capable node on the LSP MAY then use the value in the Scratch Pad field to perform time correction, if there is no follow-up message. For example, the egress node may be a PTP Boundary Clock synchronized to a Master Clock and will use the value in the Scratch Pad field to update PTP's correctionField.

6. Applicable PTP Scenarios

The proposed approach can be directly integrated in a PTP network based on the IEEE 1588 delay request-response mechanism. The RTM capable node nodes act as end-to-end transparent clocks, and typically boundary clocks, at the edges of the MPLS network, use the value in the Scratch Pad field to update the correctionField of the corresponding PTP event packet prior to performing the usual PTP processing.

7. One-step Clock and Two-step Clock Modes

One-step mode refers to the mode of operation where an egress interface updates the correctionField value of an original event message. Two-step mode refers to the mode of operation where this update is made in a subsequent follow-up message.

Processing of the follow-up message, if present, requires the downstream end-point to wait for the arrival of the follow-up message in order to combine correctionField values from both the original (event) message and the subsequent (follow-up) message. In a similar fashion, each 2-step node needs to wait for the related follow-up message, if there is one, in order to update that follow-up message (as opposed to creating a new one. Hence the first node that uses 2-step mode MUST do two things:

1. Mark the original event message to indicate that a follow-up message will be forthcoming (this is necessary in order to

Let any subsequent 2-step node know that there is already a follow-up message, and

Let the end-point know to wait for a follow-up message;

2. Create a follow-up message in which to put the RTM determined as an initial correctionField value.

IEEE 1588v2 [[IEEE.1588.2014](#)] defines this behavior for PTP messages.

Thus, for example, with reference to the PTP protocol, the PTPTType field identifies whether the message is a Sync message, Follow_up message, Delay_Req message, or Delay_Resp message. The 10 octet long Port ID field contains the identity of the source port, that is, the specific PTP port of the boundary clock connected to the MPLS network. The Sequence ID is the sequence ID of the PTP message carried in the Value field of the message.

PTP messages also include a bit that indicates whether or not a follow-up message will be coming. This bit, once it is set by a 2-step mode device, MUST stay set accordingly until the original and follow-up messages are combined by an end-point (such as a Boundary Clock).

Thus, an RTM packet, containing residence time information relating to an earlier packet, also contains information identifying that earlier packet.

For compatibility with PTP, RTM (when used for PTP packets) must behave in a similar fashion. To do this, a 2-step RTM capable egress interface will need to examine the S-bit in the Flags field of the PTP sub-TLV (for RTM messages that indicate they are for PTP) and - if it is clear (set to zero), it MUST set it and create a follow-up PTP Type RTM message. If the S bit is already set, then the RTM capable node MUST wait for the RTM message with the PTP type of follow-up and matching originator and sequence number to make the corresponding residence time update to the Scratch Pad field.

In practice an RTM operating according to two-step clock behaves like a two-steps transparent clock.

A 1-step capable RTM node MAY elect to operate in either 1-step mode (by making an update to the Scratch Pad field of the RTM message containing the PTP even message), or in 2-step mode (by making an update to the Scratch Pad of a follow-up message when its presence is indicated), but MUST NOT do both.

Two main subcases can be identified for an RTM node operating as a two-step clock:

A) If any of the previous RTM capable node or the previous PTP clock (e.g. the BC connected to the first node), is a two-step clock, the residence time is added to the RTM packet that has been created to include the associated PTP packet (i.e. follow-up message in the downstream direction), if the local RTM-capable node is also operating as a two-step clock. This RTM packet carries the related accumulated residence time and the appropriate values of the Sequence Id and Port Id (the same identifiers carried in the packet processed) and the Two-step Flag set to 1.

Note that the fact that an upstream RTM-capable node operating in the two-step mode has created a follow-up message does not require any subsequent RTM capable node to also operate in the 2-step mode, as long as that RTM-capable node forwards the follow-up message on the same LSP on which it forwards the corresponding previous message.

A one-step capable RTM node MAY elect to update the RTM follow-up message as if it were operating in two-step mode, however, it MUST NOT update both messages.

A PTP event packet (sync) is carried in the RTM packet in order for an RTM node to identify that residence time measurement must be performed on that specific packet.

To handle the residence time of the Delay request message on the upstream direction, an RTM packet must be created to carry the residence time on the associated downstream Delay Resp message.

The last RTM node of the MPLS network in addition to update the correctionField of the associated PTP packet, must also properly handle the two-step flag of the PTP packets.

B) When the PTP network connected to the MPLS and RTM node, operates in one-step clock mode, the associated RTM packet must be created by the RTM node itself. The associated RTM packet including the PTP event packet needs now to indicate that a follow up message will be coming.

The last RTM node of the LSP, if it receives an RTM message with a PTP payload indicating a follow-up message will be forthcoming, must generate a follow-up message and properly set the two-step flag of the PTP packets.

8. IANA Considerations

8.1. New RTM G-ACh

IANA is requested to reserve a new G-ACh as follows:

Value	Description	Reference
TBA1	Residence Time Measurement	This document

Table 1: New Residence Time Measurement

8.2. New RTM TLV Registry

IANA is requested to create sub-registry in Generic Associated Channel (G-ACh) Parameters Registry called "MPLS RTM TLV Registry". All code points in the range 0 through 127 in this registry shall be allocated according to the "IETF Review" procedure as specified in [\[RFC5226\]](#). Remaining code points are allocated according to the table below. This document defines the following new values RTM TLV types:

Value	Description	Reference
0	Reserved	This document
1	No payload	This document
2	PTPv2, Ethernet encapsulation	This document
3	PTPv2, IPv4 Encapsulation	This document
4	PTPv2, IPv6 Encapsulation	This document
5	NTP	This document
6-127	Reserved	IETF Consensus
128 - 191	Reserved	First Come First Served
192 - 255	Reserved	Private Use

Table 2: RTM TLV Type

8.3. New RTM Sub-TLV Registry

IANA is requested to create sub-registry in MPLS RTM TLV Registry, requested in [Section 8.2](#), called "MPLS RTM Sub-TLV Registry". All code points in the range 0 through 127 in this registry shall be allocated according to the "IETF Review" procedure as specified in [\[RFC5226\]](#). Remaining code points are allocated according to the table below. This document defines the following new values RTM sub-TLV types:

Value	Description	Reference
0	Reserved	This document
1	PTP 2-step	This document
2-127	Reserved	IETF Consensus
128 - 191	Reserved	First Come First Served
192 - 255	Reserved	Private Use

Table 3: RTM Sub-TLV Type

8.4. RTM Capability sub-TLV in OSPFv2

IANA is requested to assign a new type for RTM Capability sub-TLV from OSPFv2 Extended Link TLV Sub-TLVs registry as follows:

Value	Description	Reference
TBA2	RTM Capability	This document

Table 4: RTM Capability sub-TLV

8.5. RTM Capability sub-TLV in OSPFv3

IANA is requested to assign a new type for RTM Capability sub-TLV from future OSPFv3 Extended-LSA Sub-TLVs registry that would be part of OSPFv3 IANA registry as follows:

Value	Description	Reference
TBA3	RTM Capability	This document

Table 5: RTM Capability sub-TLV

8.6. IS-IS RTM Application ID

IANA is requested to assign a new Application ID for RTM from the Application Identifiers for TLV 251 registry as follows:

Value	Description	Reference
TBA4	RTM	This document

Table 6: IS-IS RTM Application ID

8.7. RTM_SET Sub-object RSVP Type and sub-TLVs

IANA is requested to assign a new Type for RTM_SET sub-object from Attributes TLV Space sub-registry as follows:

Typ	Name	Allowed on LSP_ATTRIBUTES	Allowed on LSP_REQUIRED_ATTRIBUTES	Allowed on LSP_Hop Attributes	Reference
TBA	RTM_SET	Yes	No	No	This document
5	sub-object				

Table 7: RTM_SET Sub-object Type

IANA requested to create new sub-registry for sub-TLV types of RTM_SET sub-object as follows:

Value	Description	Reference
0	Reserved	
1	IPv4 address	This document
2	IPv6 address	This document
3	Unnumbered interface	This document
4-127	Reserved	IETF Consensus
128 - 191	Reserved	First Come First Served
192 - 255	Reserved	Private Use

Table 8: RTM_SET object sub-object types

8.8. RTM_SET Attribute Flag

IANA is requested to assign new flag from Attribute Flags registry

Bit	Name	Attribute Flags	Attribute Flags Resv	RRO	ERO	Reference
No		Path				
TBA	RTM_SET	Yes	Yes	No	No	This document
6	T					

Table 9: RTM_SET Attribute Flag

8.9. New Error Codes

IANA is requested to assign new Error Codes from Error Codes and Globally-Defined Error Value Sub-Codes registry

Error Code	Meaning	Reference
TBA7	Duplicate TLV	This document
TBA8	Duplicate sub-TLV	This document
TBA9	RTM_SET TLV Absent	This document

Table 10: New Error Codes

9. Security Considerations

Routers that support Residence Time Measurement are subject to the same security considerations as defined in [\[RFC5586\]](#) .

In addition - particularly as applied to use related to PTP - there is a presumed trust model that depends on the existence of a trusted relationship of at least all PTP-aware nodes on the path traversed by PTP messages. This is necessary as these nodes are expected to correctly modify specific content of the data in PTP messages and proper operation of the protocol depends on this ability.

As a result, the content of the PTP-related data in RTM messages that will be modified by intermediate nodes cannot be authenticated, and the additional information that must be accessible for proper operation of PTP 1-step and 2-step modes MUST be accessible to intermediate nodes (i.e. - MUST NOT be encrypted in a manner that makes this data inaccessible).

While it is possible for a supposed compromised node to intercept and modify the G-ACh content, this is an issue that exists for nodes in general - for any and all data that may be carried over an LSP - and is therefore the basis for an additional presumed trust model associated with existing LSPs and nodes.

The ability for potentially authenticating and/or encrypting RTM and PTP data that is not needed by intermediate RTM/PTP-capable nodes is for further study.

Security requirements of time protocols are provided in [RFC 7384](#) [\[RFC7384\]](#).

10. Acknowledgements

Authors want to thank Loa Andersson, Lou Berger and Acee Lindem for their thorough reviews, thoughtful comments and, most of, patience.

11. References

11.1. Normative References

- [I-D.ietf-ospf-ospfv3-lsa-extend]
Lindem, A., Mirtorabi, S., Roy, A., and F. Baker, "OSPFv3 LSA Extendibility", [draft-ietf-ospf-ospfv3-lsa-extend-09](#) (work in progress), November 2015.
- [IEEE.1588.2014]
"Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems", IEEE Standard 1588, August 2014.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", [BCP 14](#), [RFC 2119](#), DOI 10.17487/RFC2119, March 1997, <<http://www.rfc-editor.org/info/rfc2119>>.
- [RFC3209] Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V., and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels", [RFC 3209](#), DOI 10.17487/RFC3209, December 2001, <<http://www.rfc-editor.org/info/rfc3209>>.
- [RFC3477] Kompella, K. and Y. Rekhter, "Signalling Unnumbered Links in Resource ReSerVation Protocol - Traffic Engineering (RSVP-TE)", [RFC 3477](#), DOI 10.17487/RFC3477, January 2003, <<http://www.rfc-editor.org/info/rfc3477>>.
- [RFC4385] Bryant, S., Swallow, G., Martini, L., and D. McPherson, "Pseudowire Emulation Edge-to-Edge (PWE3) Control Word for Use over an MPLS PSN", [RFC 4385](#), DOI 10.17487/RFC4385, February 2006, <<http://www.rfc-editor.org/info/rfc4385>>.
- [RFC5085] Nadeau, T., Ed. and C. Pignataro, Ed., "Pseudowire Virtual Circuit Connectivity Verification (VCCV): A Control Channel for Pseudowires", [RFC 5085](#), DOI 10.17487/RFC5085, December 2007, <<http://www.rfc-editor.org/info/rfc5085>>.

- [RFC5420] Farrel, A., Ed., Papadimitriou, D., Vasseur, JP., and A. Ayyangarps, "Encoding of Attributes for MPLS LSP Establishment Using Resource Reservation Protocol Traffic Engineering (RSVP-TE)", [RFC 5420](#), DOI 10.17487/RFC5420, February 2009, <<http://www.rfc-editor.org/info/rfc5420>>.
- [RFC5586] Bocci, M., Ed., Vigoureux, M., Ed., and S. Bryant, Ed., "MPLS Generic Associated Channel", [RFC 5586](#), DOI 10.17487/RFC5586, June 2009, <<http://www.rfc-editor.org/info/rfc5586>>.
- [RFC5905] Mills, D., Martin, J., Ed., Burbank, J., and W. Kasch, "Network Time Protocol Version 4: Protocol and Algorithms Specification", [RFC 5905](#), DOI 10.17487/RFC5905, June 2010, <<http://www.rfc-editor.org/info/rfc5905>>.
- [RFC6423] Li, H., Martini, L., He, J., and F. Huang, "Using the Generic Associated Channel Label for Pseudowire in the MPLS Transport Profile (MPLS-TP)", [RFC 6423](#), DOI 10.17487/RFC6423, November 2011, <<http://www.rfc-editor.org/info/rfc6423>>.
- [RFC6823] Ginsberg, L., Previdi, S., and M. Shand, "Advertising Generic Information in IS-IS", [RFC 6823](#), DOI 10.17487/RFC6823, December 2012, <<http://www.rfc-editor.org/info/rfc6823>>.
- [RFC7684] Psenak, P., Gredler, H., Shakir, R., Henderickx, W., Tantsura, J., and A. Lindem, "OSPFv2 Prefix/Link Attribute Advertisement", [RFC 7684](#), DOI 10.17487/RFC7684, November 2015, <<http://www.rfc-editor.org/info/rfc7684>>.

11.2. Informative References

- [I-D.ietf-tictoc-1588overmpls]
Davari, S., Oren, A., Bhatia, M., Roberts, P., and L. Montini, "Transporting Timing messages over MPLS Networks", [draft-ietf-tictoc-1588overmpls-07](#) (work in progress), October 2015.
- [RFC4202] Kompella, K., Ed. and Y. Rekhter, Ed., "Routing Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", [RFC 4202](#), DOI 10.17487/RFC4202, October 2005, <<http://www.rfc-editor.org/info/rfc4202>>.

- [RFC5226] Narten, T. and H. Alvestrand, "Guidelines for Writing an IANA Considerations Section in RFCs", [BCP 26](#), [RFC 5226](#), DOI 10.17487/RFC5226, May 2008, <<http://www.rfc-editor.org/info/rfc5226>>.
- [RFC6374] Frost, D. and S. Bryant, "Packet Loss and Delay Measurement for MPLS Networks", [RFC 6374](#), DOI 10.17487/RFC6374, September 2011, <<http://www.rfc-editor.org/info/rfc6374>>.
- [RFC7384] Mizrahi, T., "Security Requirements of Time Protocols in Packet Switched Networks", [RFC 7384](#), DOI 10.17487/RFC7384, October 2014, <<http://www.rfc-editor.org/info/rfc7384>>.

Authors' Addresses

Greg Mirsky
Ericsson

Email: gregory.mirsky@ericsson.com

Stefano Ruffini
Ericsson

Email: stefano.ruffini@ericsson.com

Eric Gray
Ericsson

Email: eric.gray@ericsson.com

John Drake
Juniper Networks

Email: jdrake@juniper.net

Stewart Bryant
Cisco Systems

Email: stbryant@cisco.com

Alexander Vainshtein
ECI Telecom

Email: Alexander.Vainshtein@ecitele.com