

Network Working Group  
Internet-Draft  
Intended status: Standards Track  
Expires: April 23, 2012

X. Fu  
M. Betts  
Q. Wang  
ZTE  
D. McDysan  
A. Malis  
Verizon  
V. Manral  
Hewlett-Packard Corp.  
October 21, 2011

RSVP-TE extensions for services aware MPLS  
draft-fuxh-mpls-delay-loss-rsvp-te-ext-00

## Abstract

With more and more enterprises using cloud based services, the distances between the user and the applications are growing. For multiple applications such as High Performance Computing and Electronic Financial markets, the response times are critical as is packet loss, while other applications require more throughput. For example, financial or trading companies are very focused on end-to-end private pipe line latency optimizations that improve things 2-3 ms. Latency and latency SLA is one of the key parameters that these "high value" customers use to select a private pipe line provider. This document extends RSVP-TE protocol to promote SLA experience of latency and packet loss application.

## 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 April 23, 2012.

## Copyright Notice

Internet-Draft

RSVP-TE for services aware MPLS

October 2011

Copyright (c) 2011 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to [BCP 78](http://trustee.ietf.org/license-info) 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

<a href="#">1.</a>	<a href="#">Introduction . . . . .</a>	<a href="#">3</a>
<a href="#">1.1.</a>	<a href="#">Conventions Used in This Document . . . . .</a>	<a href="#">3</a>
<a href="#">2.</a>	<a href="#">Performance Accumulation and Verification . . . . .</a>	<a href="#">3</a>
<a href="#">2.1.</a>	<a href="#">Signaling Extensions . . . . .</a>	<a href="#">4</a>
<a href="#">2.1.1.</a>	<a href="#">Latency Accumulation Object . . . . .</a>	<a href="#">4</a>
<a href="#">2.1.1.1.</a>	<a href="#">Latency Accumulation sub-TLV . . . . .</a>	<a href="#">5</a>
<a href="#">2.1.2.</a>	<a href="#">Required Latency Object . . . . .</a>	<a href="#">6</a>
<a href="#">2.1.3.</a>	<a href="#">Signaling Procedures . . . . .</a>	<a href="#">7</a>
<a href="#">3.</a>	<a href="#">Performance SLA Parameters Conveying . . . . .</a>	<a href="#">8</a>
<a href="#">3.1.</a>	<a href="#">Signaling Extensions . . . . .</a>	<a href="#">9</a>
<a href="#">3.1.1.</a>	<a href="#">Latency SLA Parameters subobject . . . . .</a>	<a href="#">9</a>
<a href="#">3.1.2.</a>	<a href="#">Signaling Procedure . . . . .</a>	<a href="#">12</a>
<a href="#">4.</a>	<a href="#">Security Considerations . . . . .</a>	<a href="#">12</a>
<a href="#">5.</a>	<a href="#">IANA Considerations . . . . .</a>	<a href="#">12</a>
<a href="#">6.</a>	<a href="#">References . . . . .</a>	<a href="#">13</a>
<a href="#">6.1.</a>	<a href="#">Normative References . . . . .</a>	<a href="#">13</a>
<a href="#">6.2.</a>	<a href="#">Informative References . . . . .</a>	<a href="#">13</a>
	<a href="#">Authors' Addresses . . . . .</a>	<a href="#">13</a>

## 1. Introduction

End-to-end service optimization based on latency is a key requirement for service provider. It needs to communicate latency of links and nodes including latency and latency variation as a traffic engineering performance metric is a very important requirement. [[LATENCY-REQ](#)] describes the requirement of latency traffic engineering application.

This document extend RSVP-TE to accumulate (e.g., sum) latency information of links and nodes along one LSP across multi-domain (e.g., Inter-AS, Inter-Area or Multi-Layer) so that an latency verification can be made at end points and improve the user's experience. One-way and round-trip latency collection along the LSP by signaling protocol can be supported. So the end points of this LSP can verify whether the total amount of latency could meet the latency agreement between operator and his user. When RSVP-TE signaling is used, the source can determine if the latency requirement is met much more rapidly than performing the actual end-to-end latency measurement.

One end-to-end LSP may be across some Composite Links [[CL-REQ](#)]. Even if the transport technology (e.g., OTN) implementing the component links is identical, the latency characteristics of the component links may differ. RSVP-TE message needs to carry a indication for the selection of component links based on the latecny constraint. When one end-to-end LSP traverse a server layer, there will be some latency constraint requirement for server layer. RSVP-TE message also needs to carry a indication for the FA selection or FA-LSP creation. This document extends RSVP-TE to indicate that a component links, FA or FA-LSP should meet the minimum and maximum latency value or maximum acceptable latency variation value.

Packet Loss constraint will be taken up in a future version of the draft.

## [1.1.](#) Conventions Used in This Document

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.](#) Performance Accumulation and Verification

Latency accumulation and verification applies where the full path of an multi-domain (e.g., Inter-AS, Inter-Area or Multi-Layer) TE LSP can't be or is not determined at the ingress node of the TE LSP.

Fu, et al.

Expires April 23, 2012

[Page 3]

---

Internet-Draft

RSVP-TE for services aware MPLS

October 2011

This is most likely to arise owing to TE visibility limitations. If all domains support to communicate latency as a traffic engineering metric parameter, one end-to-end optimized path with delay constraint (e.g., less than 10 ms) which satisfies latency SLAs parameter could be computed by BRPC [[RFC5441](#)] in PCE. Otherwise, it could use the mechanism defined in this section to accumulate the latency of each links and nodes along the path which is across multi-domain.

Latency accumulation and verification also applies where not all domains could support the communication latency as a traffic engineering metric parameter. The required latency could be signaled by RSVP-TE (i.e., Path and Resv message). Intermediate nodes could reject the request (Path or Resv message) if the accumulated latency is not achievable. This is essential in multiple AS use cases, but may not be needed in a single IGP level/area if the IGP is extended to convey latency information.

Node latency for a WAN could be ignored or even an average, however that was not true for the LAN cases. Whether the node latency should be accumulated or not depends on the implementation.

One domain may need to know that other domains support latency accumulation. It could be discovered in some automatic way. PCEs in different domains may play a role here. It is for further study.

### [2.1.](#) Signaling Extensions

#### [2.1.1.](#) Latency Accumulation Object

An Latency Accumulation Object is defined in this document to support the accumulation and verification of the latency. This object which can be carried in a Path/Resv message may includes two sub-TLVs. Latency Accumulation Object has the following format.

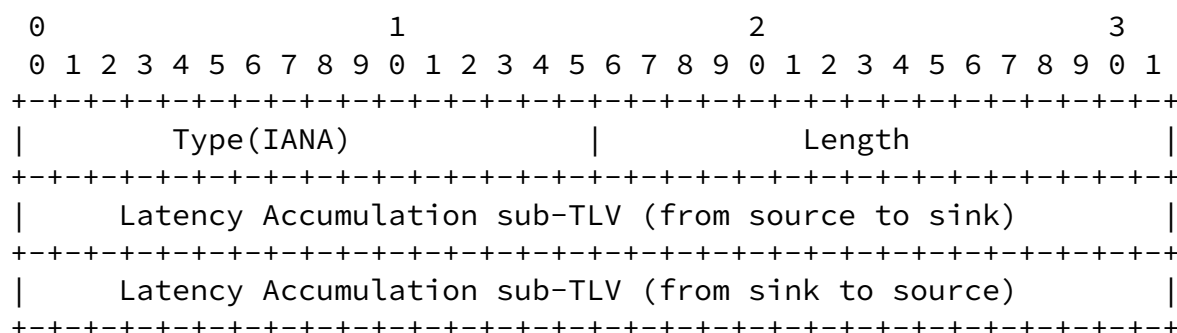


Figure 1: Format of Accumulated Latency Object

- o Latency Accumulation sub-TLV (from source to sink):It is used to accumulate the latency from source to sink along the unidirectional or bidirectional LSP. A Path message for unidirectional and bidirectional LSP must includes this sub-TLV. When sink node receives the Path message including this sub-TLV, it must copy this sub-TLV into Resv message. So the source node can receive the latency accumulated value (i.e., sum) from itself to sink node which can be used for latency verification.
- o Latency Accumulation sub-TLV (from sink to source):It is used to accumulate the latency from sink to source along the bidirectional LSP. A Resv message for the bidirectional LSP must includes this sub-TLV. So the source node can get the latency accumulated value (i.e., sum) of round-trip which can be used for latency verification. In the case of unidirectional LSP, this sub-TLV is absent.

#### [2.1.1.1](#). Latency Accumulation sub-TLV

The Sub-TLV format is defined in the next picture.



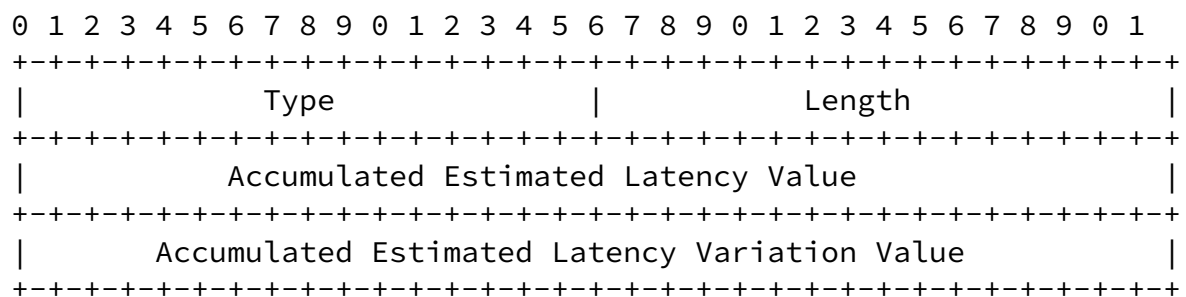


Figure 2: Format of Latency Accumulation sub-TLV

- o Type: sub-TLV type
  - \* 0: It indicates the sub-TLV is for the latency accumulation from source to sink node along the LSP.
  - \* 1: It indicates the sub-TLV is for the latency accumulation from sink to source node along the LSP.
- o Length: length of the sub-TLV value in bytes.
- o Accumulated Estimated Latency Value: a value indicates the sum of each links and nodes' latency along one direction of LSP. It MUST be quantified in units of micro-seconds and encoded as an float point value.

- o Accumulated Estimated Latency Variation Value: a value indicates the sum of each links and nodes' latency variation along one direction of LSP. Since latency variation is accumulated non-linearly. Latency variation accumulation should be in a lower priority. It MUST be quantified in units of nano-seconds and encoded as an float point value.

### 2.1.2. Required Latency Object

A required latency could be signaled by RSVP-TE message (i.e., Path and Resv). This object is carried in the LSP\_ATTRIBUTES object of Path/Resv message, object that is defined in [[RFC5420](#)]. Intermediate nodes could reject the request (Path or Resv message) if the accumulated latency value exceeds required latency value in the Required Latency Object.

If the accumulated latency is not achievable, there is no necessary to accumulate the latency for remaining domain or nodes. In order to balance the load across network links more efficiently if the absolute minimum latency is not required, intermediate nodes could choose a cost-effective path if the requested latency could easily be met. Note that this would apply inter-AS if the IGP is extended to advertise latency.

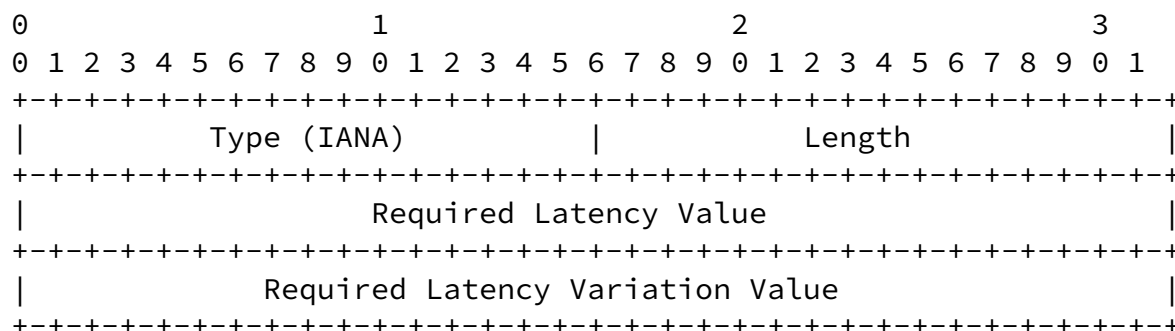


Figure 3: Required Latency Object

- o Required Latency Value: The accumulated estimated latency value should not exceed this value. It MUST be quantified in units of micro-seconds and encoded as an float point value.
- o Required Latency Variation Value: The accumulated estimated latency variation value should not exceed this value. It MUST be quantified in units of micro-seconds and encoded as an float point value.

### 2.1.3. Signaling Procedures

When the source node desires to accumulate (i.e., sum) the total latency of one end-to-end LSP, the "Latency Accumulating desired" flag (value TBD) should be set in the LSP\_ATTRIBUTES object of Path/Resv message, object that is defined in [\[RFC5420\]](#). If the source node makes the intermediate node have the capability to verify the accumulated latency, the "Latency Verifying desired" flag (value TBD)

should be also set in the LSP\_ATTRIBUTES object of Path/Resv message.

A source node initiates latency accumulation for a given LSP by adding Latency Accumulation object to the Path message. The Latency Accumulation object only includes one sub-TLV (sub-TLV type=0) where it is going to accumulate the latency value of each links and nodes along path from source to sink. If latency verifying is desired, the source node also adds the Required Latency Object to the Path message.

When the downstream node receives Path message and if the "Latency Accumulating desired" is set in the LSP\_ATTRIBUTES, it accumulates the latency of link and node based on the accumulated latency value of the sub-TLV (sub-TLV type=0) in Latency Accumulation object before it sends Path message to downstream.

If the "Latency Verifying desired" is set in the LSP\_ATTRIBUTES, downstream node will check whether the Accumulated Estimated Latency and Variation value exceeds the Required Latency and Variation value. If the accumulated latency is not achievable, there is no necessary to accumulate the latency for remaining domain or nodes. It MUST generate a error message with a "Accumulated Latency couldn't meet the required latency" indication (TBD by IANA).

If the intermediate node (e.g., entry node of one domain) couldn't support the latency accumulation function, it MUST generate a error message with a "Latency Accumulation unsupported" indication (TBD by IANA).

If the intermediate node (e.g., entry node of one domain) couldn't support the latency verify function, it MUST generate a error message with a "Latency Verify unsupported" indication (TBD by IANA).

When the sink node of LSP receives the Path message and the "Latency Accumulating desired" is set in the LSP\_ATTRIBUTES, it copy the Accumulated Estimated Latency and Variation value in the Latency Accumulation sub-TLV (sub-TLV type=0) of Path message into the one of Resv message which will be forwarded hop by hop in the upstream direction until it arrives the source node. Then source node can get the latency sum value from source to sink for unidirectional and



If the LSP is a bidirectional one and the "Latency Accumulating desired" is set in the LSP\_ATTRIBUTES, it adds another Latency Accumulation sub-TLV (sub-TLV type=1) into the Latency Accumulation object of Resv message where latency of each links and nodes along path will be accumulated from sink to source into this sub-TLV.

If the LSP is a bidirectional one and the "Latency Verifying desired" is set in the LSP\_ATTRIBUTES, it copy the Required Latency and Variation value in the Required Latency Object of Path message into the Resv message.

When the upstream node receives Resv message and if the "Latency Accumulating desired" is set in the LSP\_ATTRIBUTES, it accumulates the latency of link and node based on the latency value in sub-TLV (sub-TLV type=1) before it continues to sends Resv message.

If the "Latency Verifying desired" is set in the LSP\_ATTRIBUTES, it will check whether the latency sum of Accumulated Estimated Latency and Variation value in each Latency Accumulation sub-TLV exceeds the Required Latency and Variation value. If the accumulated latency is not achievable, there is no necessary to accumulate the latency for remaining domain or nodes. It MUST generate a error message with a "Accumulated Latency couldn't meet the required latency" indication (TBD by IANA).

After source node receive Resv message, it can get the total latency value of one way or round-trip from Latency Accumulation object. So it can confirm whether the latency value meet the latency SLA or not.

### [3.](#) Performance SLA Parameters Conveying

[CL-REQ] introduces Composite Link into MPLS network. In order to assign the LSP to one of component links with different latency characteristics, RSVP-TE message MUST convey latency SLA parameter to the end points of Composite Links where it can select one of component links or trigger the creation of lower layer connection which MUST meet latency SLA parameter.

One end-to-end LSP (e.g., in IP/MPLS or MPLS-TP network) may traverse a FA-LSP of server layer (e.g., OTN rings). There will be some latency constraint requirement for server layer. RSVP-TE message also needs to carry a indication for the FA selection or FA-LSP creation.

The RSVP-TE message needs to carry a indication of request minimum

latency, maximum acceptable latency value and maximum acceptable delay variation value. The end point will take these parameters into account for selection or creation of component link, FA selection or FA-LSP.

### [3.1.](#) Signaling Extensions

This document defines extensions to and describes the use of RSVP-TE [[RFC3209](#)], [[RFC3471](#)], [[RFC3473](#)] to explicitly convey the latency SLA parameter for the selection or creation of component link or FA/FA-LSP. Specifically, in this document, Latency SLA Parameters TLV are defined and added into ERO as a subobject.

#### [3.1.1.](#) Latency SLA Parameters subobject

A new OPTIONAL subobject of the EXPLICIT\_ROUTE Object (ERO) is used to specify the latency SLA parameters including a indication of request minimum latency, request maximum acceptable latency value and request maximum acceptable latency variation value. It can be used for the following scenarios.

- o One end-to-end LSP may traverse a server layer FA-LSP. This subobject of ERO can indicate that FA selection or FA-LSP creation shall be based on this latency constraint. The boundary nodes of multi-layer will take these parameters into account for FA selection or FA-LSP creation.
- o One end-to-end LSP may be across some Composite Links [[CL-REQ](#)]. This subobject of ERO can indicate that a traffic flow shall select a component link with some latency constraint values as specified in this subobject.

This Latency SLA Parameters ERO subobject has the following format. It follows a subobject containing the IP address, or the link identifier [[RFC3477](#)], associated with the TE link on which it is to be used.

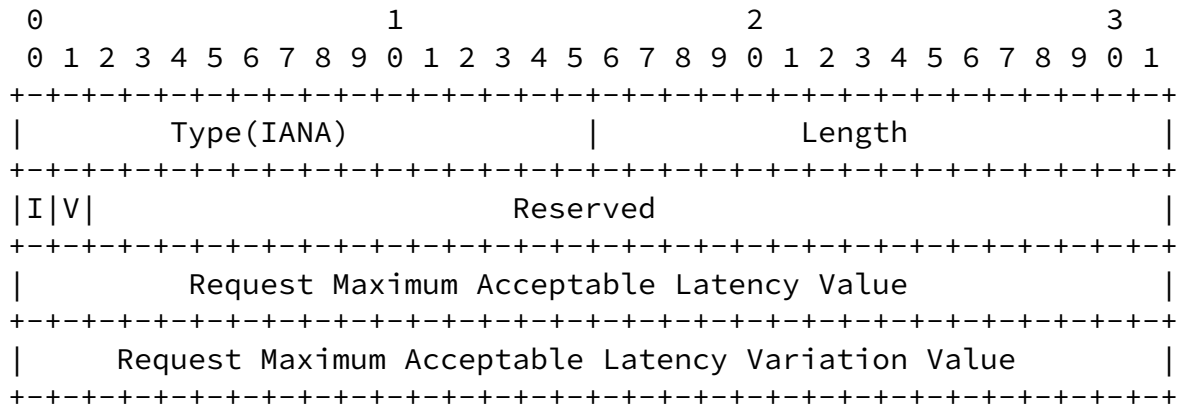


Figure 4: Format of Latency SLA Parameters TLV

- o I bit: a one bit field indicates whether a traffic flow shall select a component link with the minimum latency value or not. It can also indicate whether one end-to-end LSP shall select a FA or trigger a FA-LSP creation with the minimum latency value or not when it traverse a server layer.
- o V bit: a one bit field indicates whether a traffic flow shall select a component link with the minimum latency variation value or not. It can also indicate whether one end-to-end LSP shall select a FA or trigger a FA-LSP creation with the minimum latency variation value or not when it traverse a server layer.
- o Request Maximum Acceptable Latency Value: a value indicates that a traffic flow shall select a component link with a maximum acceptable latency value. It can also indicate one end-to-end LSP shall select a FA or trigger a FA-LSP creation with a maximum acceptable latency value when it traverse a server layer. It MUST be quantified in units of micro-seconds and encoded as an float point value.
- o Request Maximum Acceptable Latency Variation Value: a value indicates that a traffic flow shall select a component link with a maximum acceptable latency variation value. It can also indicate one end-to-end LSP shall select a FA or trigger a FA-LSP creation with a maximum acceptable latency variation value when it traverse

a server layer. It MUST be quantified in units of nano-seconds and encoded as an float point value.

Following is an example about how to use these parameters. Assume there are following component links within one composite link.

- o Component link1: latency = 50 ms, latency variation = 15 ns

- o Component link2: latency = 100 ms, latency variation = 6 ns
- o Component link3: latency = 200 ms, latency variation = 3 ns
- o Component link4: latency = 300 ms, latency variation = 1 ns

Assume there are following request information.

- o Request minimum latency = FALSE
- o Request minimum latency variation= FALSE
- o Maximum Acceptable Latency Value= 150 ms
- o Maximum Acceptable Latency Variation Value = 10 ns

Only Component link2 could be qualified.

- o Request minimum latency = FALSE
- o Request minimum latency variation= FALSE
- o Maximum Acceptable Latency Value= 350 ms
- o Maximum Acceptable Latency Variation Value = 10 ns

Component link2/3/4 could be qualified. Which component link is selected depends on local policy.

- o Request minimum latency = FALSE
- o Request minimum latency variation= TRUE
- o Maximum Acceptable Latency Value= 350 ms
- o Maximum Acceptable Latency Variation Value = 10 ns

Only Component link4 could be qualified.

- o Request minimum latency = TRUE
- o Request minimum latency variation= FALSE

- o Maximum Acceptable Latency Value= 350 ms
- o Maximum Acceptable Latency Variation Value = 10 ns

Only Component link2 could be qualified.

Request minimum latency = TRUE

Request minimum latency variation= TRUE

Maximum Acceptable Latency Value= 350 ms

Maximum Acceptable Latency Variation Value = 10 ns

In this case, there is no any qualified component links. But priority may be used for latency and variation, so one of component links could be still selected.

### [3.1.2.](#) Signaling Procedure

When a intermediate node receives a PATH message containing ERO and finds that there is a Latency SLA Parameters ERO subobject immediately behind the IP address or link address sub-object related to itself, if the node determines that it's a region edge node of FA-LSP or an end point of a composite link [[CL-REQ](#)], then, this node

extracts latency SLA parameters (i.e., request minimum, request maximum acceptable and request maximum acceptable latency variation value) from Latency SLA Parameters ERO subobject. This node used these latency parameters for FA selection, FA-LSP creation or component link selection.

If the intermediate node couldn't support the latency SLA, it MUST generate a PathErr message with a "Latency SLA unsupported" indication (TBD by IANA). If the intermediate node couldn't select a FA or component link, or create a FA-LSP which meet the latency constraint defined in Latency SLA Parameters ERO subobject, it must generate a PathErr message with a "Latency SLA parameters couldn't be met" indication (TBD by IANA).

#### [4.](#) Security Considerations

This document raises no new security issues.

#### [5.](#) IANA Considerations

TBD

Fu, et al.

Expires April 23, 2012

[Page 12]

---

Internet-Draft

RSVP-TE for services aware MPLS

October 2011

#### [6.](#) References

##### [6.1.](#) Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", [BCP 14](#), [RFC 2119](#), March 1997.
- [RFC3209] Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V., and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels", [RFC 3209](#), December 2001.
- [RFC3473] Berger, L., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions", [RFC 3473](#), January 2003.
- [RFC3477] Kompella, K. and Y. Rekhter, "Signalling Unnumbered Links in Resource ReSerVation Protocol - Traffic Engineering (RSVP-TE)", [RFC 3477](#), January 2003.

- [RFC3630] Katz, D., Kompella, K., and D. Yeung, "Traffic Engineering (TE) Extensions to OSPF Version 2", [RFC 3630](#), September 2003.
- [RFC4203] Kompella, K. and Y. Rekhter, "OSPF Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", [RFC 4203](#), October 2005.

## [6.2.](#) Informative References

- [CL-REQ] C. Villamizar, "Requirements for MPLS Over a Composite Link", [draft-ietf-rtgwg-cl-requirement-02](#) .
- [EXPRESS-PATH]  
S. Giacalone, "OSPF Traffic Engineering (TE) Express Path", [draft-giacalone-ospf-te-express-path-01](#) .
- [LATENCY-REQ]  
X. Fu, "Traffic Engineering architecture for services aware MPLS", [draft-fuxh-mpls-delay-loss-te-framework-02](#) .
- [ietf-mpls-loss-delay]  
D. Frost, "Packet Loss and Delay Measurement for MPLS Networks", [draft-ietf-mpls-loss-delay-03](#) .

### Authors' Addresses

Xihua Fu  
ZTE

Email: [fu.xihua@zte.com.cn](mailto:fu.xihua@zte.com.cn)

Malcolm Betts  
ZTE

Email: malcolm.betts@zte.com.cn

Qilei Wang  
ZTE

Email: wang.qilei@zte.com.cn

Dave McDysan  
Verizon

Email: dave.mcdysan@verizon.com

Andrew Malis  
Verizon

Email: andrew.g.malis@verizon.com

Vishwas Manral  
Hewlett-Packard Corp.  
191111 Pruneridge Ave.  
Cupertino, CA 95014  
US

Phone: 408-447-1497  
Email: vishwas.manral@hp.com  
URI: