Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: September 15, 2011

X. Fu
M. Betts
Q. Wang
ZTE
D. McDysan
A. Malis
Verizon

March 14, 2011

GMPLS extensions to communicate latency as a traffic engineering performance metric

draft-wang-ccamp-latency-te-metric-03

Abstract

Latency is such requirement that must be achieved according to the Service Level Agreement (SLA) between customers and service providers. Network Performance Objective (NPO) defined in ITU-T Y.1540 and Y.1541 is used for describing the meaning and numerical values performance parameters traversing multiple packet networks. The definitions of the packet network performance parameters are often also used as the basis of SLAs service providers, but possibly with different numerical values. A SLA is a part of a service contract where the level of service is formally defined between service providers and customers. For example, the service level includes platinum, golden, silver and bronze. Different service level may associate with different protection/restoration requirement. Latency can also be associated with different service level. The user may select a private line provider based on the ability to meet a latency SLA.

The key driver for latency is stock/commodity trading applications that use data base mirroring. A few milli seconds can impact a transaction. 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. Other key applications like video gaming, conferencing and storage area networks require stringent latency and bandwidth.

This document describes the requirements and mechanisms to communicate latency as a traffic engineering performance metric in today's network which is consisting of potentially multiple layers of packet transport network and optical transport network in order to meet the latency SLA between service provider and his customers. This document also extends RSVP-TE and IGP to support these requirement. These extensions are intended to advertise and convey

the latency information of nodes and links as traffic engineering performance metric.

Status of this Memo

This Internet-Draft will expire on September 15, 2011.

Copyright Notice

Table of Contents

1. Introduction

In a network, latency, a synonym for delay, is an expression of how much time it takes for a packet/frame of data to get from one designated point to another. In some usages, latency is measured by sending a packet/frame that is returned to the sender and the round-trip time is considered the latency of bidirectional co-routed or associated LSP. One way time is considered as the latency of unidirectional LSP. The one way latency may not be half of the round-trip latency in the case that the transmit and receive directions of the path are of unequal lengths.

Latency on a connection has two sources: Node latency which is caused by the node as a result of process time in each node and: Link latency as a result of packet/frame transit time between two neighbouring nodes or a FA-LSP/Composit Link [CL-REQ]. Latency variation is a parameter that is used to indicate the variation range of the latency value. These values should be made available to the control plane and management plane prior to path computation. This allows path computation to select a path that will meet the latency SLA.

In many cases, latency is a sensitive topic. For example, two stock exchanges (e.g.,one in Chicago and another in New York) need to communicate with each other. A few ms can result in large impact on service. Some customers would pay for the latency performance. SLA contract which includes the requirement of latency is signed between service providers and customers. Service provider should assure that the network path latency MUST be limited to a value lower than the upper limit. In the future, latency optimization will be needed by more and more customers. For example, some customers pay for a private pipe line with latency constraint (e.g., less than 10 ms) which connects to Data Center. If this "provisioned" latency of this private pipe line couldn't meet the SLA, service provider may transfer customer's service to other Data Centers. Service provider may have many layers of pre-defined restoration for this transfer, but they have to duplicate restoration resources at significant cost. So service provider needs some mechanisms to avoid the duplicate restoration and reduce the network cost.

Measurement mechanism for link latency has been defined in many technologies. For example, the measurement mechanism for link latency has been provided in ITU-T [G.8021] and [Y.1731] for Ethernet. The link transit latency between two Ethernet equipments can be measured by using this mechanism. Similarly, overhead byte and measurement mechanism of latency has been provided in OTN (i.e., ITU-T [G.709]). In order to measure the link latency between two OTN nodes, PM&TCM which include Path Latency Measurement field and flag

used to indicate the beginning of measurement of latency is added to the overhead of ODUk. Node latency can also be recorded at each node by recording the process time between the beginning and the end. The measurement mechanism of links and nodes is out scope of this document.

Current operation and maintenance mode of latency measurement is high in cost and low in efficiency. The latency can only be measured after the connection has been established, if the measurement indicates that the latency SLA is not met then another path is computed, set up and measured. This "trial and error" process is very inefficient. To avoid this problem a means of making an accurate prediction of latency before a path is establish is required.

This document describes the requirements and mechanisms to communicate latency as a traffic engineering performance metric in today's network which is consisting of potentially multiple layers of packet transport network and optical transport network in order to meet the latency SLA between service provider and his customers. This document extends IGP to advertise and convey the latency attributes and latency variation as traffic engineering performance metric. Thus path computation entity can have a good knowledge of the latency traffic engineering database.

This document extends RSVP-TE protocol 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 source node. 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.

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.

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. Requirements Identification and Solution Consideration

2.1. Requirements Identification

End-to-end service optimization based on latency is a key requirement for service provider. This type of function will be adopted by their "premium" service customers. They would like to pay for this "premium" service. After these premium services are deployed, they will also expand to their own customers. Following key requirements associated with latency is identified.

o Communication latency of links and nodes including latency and latency variation as a traffic engineering performance metric is a very important requirement.

o End-to-end service optimization based on latency constraint is a key requirement for service provider. Latency on a route level will help carriers' customers to make his provider selection decision.

  * Path computation entity MUST have the capability to compute one end-to-end path with latency constraint. For example, it MUST have the capability to compute a route with x amount bandwidth and less than y ms of latency limit based on the latency traffic engineering database.

  * It should also support combined routing constraints with pre-defined priorities, e.g., SRLG diversity, latency and cost.

o 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. In order to assign the LSP to one of component links with different latency characteristics, following related requirements are from [CL-REQ].

  * The solution SHALL provide a means to indicate that a traffic flow shall select a component link with the minimum latency value.

  * The solution SHALL provide a means to indicate that a traffic flow shall select a component link with a maximum acceptable latency value as specified by protocol.

* The solution SHALL provide a means to indicate that a traffic flow shall select a component link with a maximum acceptable latency variation value as specified by protocol.

o RSPV-TE should support the accumulation (e.g., sum) of 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 validation decision can be made at the source node. One-way and round-trip latency collection along the LSP by signaling protocol and latency verification at the end of LSP should be supported.

2.2. Solution Consideration

o The latency performance metric MUST be advertised into path computation entity by IGP (etc., OSPF-TE or IS-IS-TE) to perform route computation and network planning based on latecny SLA target.

* Data plane is responsible for measuring the latency (e.g., latency and latency variation). Latency measurement can be provided by different technologies. This information will be provided to the Control Plane. In order to monitor the performance, pro-active latency measurement is required. Generally, every 15 minutes or 24 hours, the value of latency and latency variation should be collected. Similarly, on demand latency measurement is required due to the goal of maintenance. This can be done every fixed time interval (e.g., 5 minutes or 1 hour). The method used to measure the latency of links and nodes is out scope of this document.

* Control plane is responsible for advertising and collecting the latency value of links and nodes by IGP (i.e., OSPF-TE/ IS-IS-TE). Latency characteristics of these links and nodes may change dynamically. In order to control IGP messaging and avoid being unstable when the latency and latency variation value changes, a threshold and a limit on rate of change MUST be configured to control plane.

o When the Composite Links [CL-REQ] is advertised into IGP, there are following solution consideration.

* The latency of composite link may be the range (e.g., at least minimum and maximum) latency value of all component links. The latency of composite link may also be the maximum latency value of all component links. In these cases, only partial information is transmited in the IGP. So the path computation entity has insufficient information to determine whether a particular path can support its delay requirements. This leads

to signaling crankback.

* The IGP may be extended to advertise latency of each component link within one Composite Link.

o 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.

* The RSVP-TE message needs to carry a indication of request minimum latency, maximum acceptable latency value and maximum acceptable delay variation value for the component link selection or creation. The composite link will take these parameters into account when assigning traffic of LSP to a component link.

o 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). The boundary nodes of the FA-LSP SHOULD be aware of the latency information of this FA-LSP (e.g., latency and latency variation).

* If the FA-LSP is able to form a routing adjacency and/or as a TE link in the client network, the latency value of the FA-LSP can be as an input to a transformation that results in a FA traffic engineering metric and advertised into the client layer routing instances. Note that this metric will include the latency of the links and nodes that the trail traverses.

* If the latency information of the FA-LSP changes (e.g., due to a maintenance action or failure in OTN rings), the boundary node of the FA-LSP will receive the TE link information advertisement including the latency value which is already changed and if it is over than the threshold and a limit on rate of change, then it will compute the total latency value of the FA-LSP again. If the total latency value of FA-LSP changes, the client layer MUST also be notified about the latest value of FA. The client layer can then decide if it will accept the increased latency or request a new path that meets the latency requirement.

* When one end-to-end LSP traverse a server layer, there will be some latency constraint requirement for the segment route in server layer. So RSVP-TE message needs to carry a indication of request minimum latency, maximum acceptable latency value and maximum acceptable delay variation value for the FA selection or FA-LSP creation. The boundary nodes of FA-LSP

will take these parameters into account for FA selection or FA-LSP creation.

o Restoration, protection and equipment variations can impact "provisioned" latency (e.g., latency increase). The change of one end-to-end LSP latency performance MUST be known by source and/or sink node. So it can inform the higher layer network of a latency change. The latency change of links and nodes will affect one end-to-end LSP's total amount of latency. Applications can fail beyond an application-specific threshold. Some remedy mechanism could be used.

  * Some customers may insist on having the ability to re-route if the latency SLA is not being met. If a "provisioned" end-to-end LSP latency could not meet the latency agreement (e.g., latency or latency variation) between operator and his user, then re-routing could be triggered based on the local policy. Pre-defined or dynamic re-routing could be triggered to handle this case. The latency performance of pre-defined or dynamic re-routing LSP MUST meet the latency SLA parameter. In the case of predefined re-routing, the large amounts of redundant capacity may have a significant negative impact on the overall network cost. Dynamic re-routing also has to face the risk of resource limitation. So the choice of mechanism MUST be based on SLA or policy. In the case where the latency SLA cannot be met after a re-route is attempted, control plane should report an alarm to management plane. It could also try restoration for several times which could be configured.

  * As a result of the change of links and nodes latency in the LSP, current LSP may be frequently switched to a new LSP with a appropriate latency value. In order to avoid this, the solution SHOULD indicate the switchover of the LSP according to maximum acceptable change latency value.

3. Control Plane Solution

In order to meet the requirements which have been identified in section 3, this document defines following four phases.

o The first phase is the advertisement of the latency information by routing protocol (i.e., OSPF-TE/IS-IS-TE), including latency of nodes and links, a FA-LSP or Composite Link [CL-REQ] between two neighbour and latency variation, so path computation entity can be aware of the latency of nodes and links.

o In the second phase, path computation entity is responsible for end-to-end path computation with latency constraint (e.g., less than 10 ms) combining other routing constraint parameters (e.g., SRLG, cost and bandwidth). How does the path computation entity compute the latency variation of one end-to-end connection can be refered to ITU-T Y.1540.

o The third phase is to convey the latency SLA parameters for the selection or creation of component link or FA/FA-LSP. One end-to-end LSP may be across some Composite Links or server layers, so it can convey latency SLA parameters by RSVP-TE message.

o The last phase is the latency collection and verification. This stage could be optional. It could accumulate (e.g., sum) latency information along the LSP across multi-domain (e.g., Inter-AS, Inter-Area or Multi-Layer) by RSVP-TE signaling message to verify the total latency at the end of path.

3.1. Latency Advertisement

A node in the packet transport network or optical transport network can detect the latency value of link which connects to it. Also the node latency can be recorded at every node. Then latency values of TE links, Composit Links [CL-REQ] or FAs, latency values of nodes and latency variation are notified to the IGP. If any latency values change and over than the threshold and a limit on rate of change, then the change MUST be notified to the IGP again. As a result, path computation entity can have every node and link latency values and latency variation in its view of the network, and it can compute one end-to-end path with latency constraint. It needs to extend IGP protocol (i.e., OSPF-TE/IS-IS-TE).

3.1.1. Routing Extensions

Following is the extensions to OSPF-TE/IS-IS-TE to support the advertisement of the node latency value, link latency and latency variation.

3.1.1.1. OSPF-TE Extension

OSPF-TE routing protocol can be used to carry latency performance metric by adding a sub-TLV to the TE link defined in [RFC4203]. As defined in [RFC3630] and [RFC4203], the top-level TLV can take one of two values (1) Router address or (2) Link. Latency sub-TLV of link is added behind the top-level TLV. It includes estimated latency and latency variation value.

This link attribute may also take into account the latency of a

network element (node), i.e., the latency between the incoming port and the outgoing port of a network element. If the link attribute were to include node latency AND link latency, then when the latency calculation is done for paths traversing links on the same node then the node latency can be subtracted out. Following is the link Latency sub-TLV format.

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|        Type(IANA)             |           Length              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      Estimated Latency Value                  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                 Estimated Latency Variation Value             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 1: Format of the Latency sub-TLV

- o Estimated Latency Value: a value indicates the latency of link or node.

- o Estimated Latency Variation Value: a value indicates the variation range of the estimated latency value.

3.1.1.2. IS-IS-TE Extension

TBD

3.1.1.3. Routing Extensions for Bundle Link/Composite Link

[Editor Notes:Some discussion have been raised in RTGWG Mailing list.]

Some people are discussing having an IGP adjacency (and metric) for a composite link but a separate advertisement that contains parameters, such as those listed here.

3.2. Latency SLA Parameters Conveying

3.2.1. 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.

3.2.1.1. Latency SLA Parameters ERO 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.

```
   0                   1                   2                   3
   0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
  +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
  |        Type(IANA)             |           Length              |
  +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
  |I|V|                       Reserved                            |
  +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
  |         Request Maximum Acceptable Latency Value              |
  +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
  |     Request Maximum Acceptable Latency Variation Value        |
  +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 2: 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.

- 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.

- 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.

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

- Component link1: latency = 5ms, latency variation = 15 us

- Component link2: latency = 10ms, latency variation = 6 us

- Component link3: latency = 20ms, latency variation = 3 us

- Component link4: latency = 30ms, latency variation = 1 us

Assume there are following request information.

- Request minimum latency = FALSE

- Request minimum latency variation= FALSE

- Maximum Acceptable Latency Value= 15 ms

- Maximum Acceptable Latency Variation Value = 10us

Only Component link2 could be qualified.

- Request minimum latency = FALSE

- Request minimum latency variation= FALSE

- Maximum Acceptable Latency Value= 35 ms

- Maximum Acceptable Latency Variation Value = 10us

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= 35 ms
- o Maximum Acceptable Latency Variation Value = 10us

Only Component link4 could be qualified.

- o Request minimum latency = TRUE
- o Request minimum latency variation= FALSE
- o Maximum Acceptable Latency Value= 35 ms
- o Maximum Acceptable Latency Variation Value = 10us

Only Component link2 could be qualified.

Request minimum latency = TRUE

Request minimum latency variation= TRUE

Maximum Acceptable Latency Value= 35 ms

Maximum Acceptable Latency Variation Value = 10us

In this case, there is no any qualified component links.

3.2.1.2. 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 INNA). 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 INNA).

3.3. Latency 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. 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 accumulat 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.

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.

3.3.1. Signaling Extensions

3.3.1.1. 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.

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|        Type(IANA)             |           Length              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     Latency Accumulation sub-TLV (from source to sink)        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     Latency Accumulation sub-TLV (from sink to source)        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 3: 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.

3.3.1.1.1. Latency Accumulation sub-TLV

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

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|             Type              |           Length              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|           Accumulated Estimated Latency Value                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|       Accumulated Estimated Latency Variation Value           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 4: 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.

o Accumulated Estimated Latency Variation Value: a value indicates the sume of each links and nodes' latency variation along one direction of LSP. Since latecny variation is accumulated non-

linearly. Latency variation accumulatoin should be in a lower priority.

3.3.1.2. 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 exceeds require 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.

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|           Type (INNA)         |            Length             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Required Latency Value                    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|              Required Latency Variation Value                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 5: Required Latency Object

- o Required Latency Value: The accumulated estimated latency value should not exceed this value.

- o Required Latency Variation Value: The accumulated estimated latency variation value should not exceed this value.

3.3.1.3. 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 desred" 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 desred, 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 downsteam.

If the "Latency Verifying desred" 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 INNA).

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 INNA).

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 INNA).

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 bidirectional LSP.

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 one of 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 desred" 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 INNA).

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.

4. Security Considerations

TBD

5. IANA Considerations

TBD

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 .

[G.709] ITU-T Recommendation G.709, "Interfaces for the Optical Transport Network (OTN)", December 2009.

Authors' Addresses

Xihua Fu
ZTE

Email: 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