CCAMP Working Group Zafar Ali
Internet Draft George Swallow
Intended status: Standard Track Clarence Filsfils
Expires: April 18, 2014 Luyuan Fang
Cisco Systems
Kenji Kumaki
KDDI Corporation
Ruediger Kunze
Deutsche Telekom AG
Daniele Ceccarelli
Ericsson
Xian Zhang
Huawei
October 19, 2013

Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extension for Signaling Objective Function and Metric Bound

draft-ali-ccamp-rc-objective-function-metric-bound-04.txt

Status of this Memo

This Internet-Draft will expire on April 18, 2014.

Copyright Notice

Abstract

In particular networks such as those used by financial institutions, network performance criteria such as latency are becoming critical to data path selection. However cost is still an important consideration. This leads to a situation where path calculation involves multiple metrics and more complex objective functions.

When using GMPLS control plane, there are many scenarios in which a node may need to request a remote node to perform path computation or expansion, like for example multi-domain LSP setup, Generalized Multi-Protocol Label Switching (GMPLS) User-Network Interface (UNI) or simply the utilization of a loose ERO in intra domain signaling. In such cases, the node requesting for the setup of an LSP needs to convey the required objective function to the remote node, to enable it to perform route computation in the desired fashion. Similarly, there are cases the ingress needs to indicate a TE metric bound for a loose segment that is expanded by a remote node.

This document defines extensions to the RSVP-TE Protocol to allow an ingress node to request the required objective function for the route computation, as well as a metric bound to influence route computation decisions at a remote node(s).

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 RFC 2119 [RFC2119].

Table of Contents

1. Introduction

As noted in [OSPF-TE-METRIC] and [ISIS-TE-METRIC], in certain networks such as financial information networks (e.g. stock market data providers), performance criteria such as latency are becoming critical to data path selection along with other metrics. Such networks may require selection of a path that minimizes end-to-end latency. Or a path may need to be found that minimized some other TE metric(s), while subject to a latency bound. Thus there is a requirement to be able to find end-to-end paths with different optimization criteria.

When the entire route for an LSP is computed at the ingress node, this requirement can be met by a local decision at that node. However, there are scenarios where partial or full route computations are performed by remote nodes. The scenarios include (but are not limited to):

- LSPs with loose hops in the Explicit Route Object (ERO), including intra-domain LSPs.

- GMPLS-UNI where route computation may be performed by the UNI-Network (server) node [RFC4208];

. Multi domain LSP setup with per domain path computation;

In these scenarios, there is a need for the ingress node to convey the optimization criteria (e.g., IGP cost, TE cost, hop counts, latency, etc.) to be used for the path computation to the node performing route computation or expansion. Similarly, there is a need for the ingress node to indicate a TE metric bound for the loose segment being expanded by a remote node.

[RFC5541] defines extensions to the Path Computation Element communication Protocol (PCEP) to allow a Path Computation Client (PCC) indicate in a path computation request the desired objective function. [RFC5440] and [ID-SERVICE-AWARE] defines extension to the PCEP to allow a PCC indicate in a path computation request a bound on given TE metric(s). This draft defines similar mechanisms for the RSVP-TE protocol allowing an ingress node to indicate in a Path request the desired objective function along with any associated TE metric bound(s). The nodes performing route expansion use this information to find the "best" candidate route.

2. RSVP-TE signaling extensions

This section defines RSVP-TE signaling extensions required to address the above-mentioned requirements. Two new ERO subobject types, Objective Function (OF) and Metric, are defined. Their purpose is as follows.

. OF subobject conveys a set of one or more specific optimization criteria that needs be followed in expanding route of a TE-LSP in MultiProtocol Label Switching (MPLS) and GMPLS networks.

. Metric Bound subobject indicates the bound on the path metric that needs to be observed for the loose segment to be considered as acceptable by the ingress node.

The scope of the Metric and OF subobjects is the node performing the expansion for loose ERO and the subsequent ERO subobject that identifies an abstract node. The following subsection provides the details.

2.1. Objective Function (OF) Subobject

A new ERO subobject type Objective Function (OF) is defined in order for the ingress node to indicate the required objective function on a loose hop. The ERO subobject type OF is optional. It MAY be carried within an ERO object of RSVP-TE Path message

and its scope is limited to previous ERO subobject that identifies an abstract node. For more details please refer to the Processing Rules for the OF Subobjects section.

The OF subobject 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|L|    Type     |     Length    |    OF Code    |   Reserved    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

The fields of OF subobject are defined as follows:

L bit: The L bit MUST be set to represent a loose hop in the explicit route.

Type: The Type is to be assigned by IANA (suggested value: 66).

Length: The Length contains the total length of the subobject in bytes, including the Type field, the Length field. The Length of the subobject is 4.

OF Code (1 byte): The identifier of the objective function. The following OF code values are suggested. These values are to be assigneyd by IANA.

* OF code value 0 is reserved.

* OF code value 1 (to be assigned by IANA) is for Minimum TE Metric Cost Path (MTMCP) OF defined in this document. See definition of MTCP OF in the following.

* OF code value 2 (to be assigned by IANA) is for Minimum Interior Gateway Protocol (IGP) Metric Cost Path (MIMCP) OF defined in the following.

* OF code value 3 (to be assigned by IANA) is for Minimum Load Path (MLP) OF as defined in RFC5541.

* OF code value 4 (to be assigned by IANA) is for Maximum Residual Bandwidth Path (MBP) OF as defined in RFC5541.

* OF code value 5 (to be assigned by IANA) is for Minimize Aggregate Bandwidth Consumption (MBC) OF as defined in RFC5541.

* OF code value 6 (to be assigned by IANA) is for Minimize the Load of the most loaded Link (MLL) OF as defined in RFC5541.

* OF code value 7 is skipped (to keep the objective function code values consistent between [RFC5541] and this draft.

* OF code value 8 (to be assigned by IANA) is for Minimum Latency Path (MLP) OF defined in this document. See definition of MLP OF in the following.

* OF code value 9 (to be assigned by IANA) is for Minimum Latency Variation Path (MLVP) OF defined in this document. See definition of MLVP OF in the following.

Other objective functions may be defined in future.

Reserved (5 bytes): This field MUST be set to zero on transmission and MUST be ignored on receipt.

2.1.1. Minimum TE Metric Cost Path Objective Function

Minimum TE Metric Cost Path (MTMCP) OF is defined as an Objective Function where a path is computed such that the sum of the TE metric of the links along the path is minimized. In the context of loose hop expansion, the ERO expanding node MUST try to find a route such that the sum of the TE metric of the links along the route is minimized.

2.1.2. Minimum IGP Metric Cost Path Objective Function

Minimum IGP Metric Cost Path (MIMCP) OF is defined as an Objective Function where a path is computed such that the sum of the IGP metric of the links along the path is minimized. In the context of loose hop expansion, the ERO expanding node MUST try to find a route such that the sum of the IGP metric of the links along the route is minimized.

2.1.3. Minimum Latency Path Objective Function

Minimum Latency Path (MLP) OF is defined as an Objective Function where a path is computed such that latency of the path is minimized. In the context of loose hop expansion, the ERO expanding node MUST try to find a route such that overall latency of the loose hop is minimized.

2.1.4. Minimum Latency Variation Path Objective Function

Minimum Latency Variation Path (MLVP) OF is defined as an Objective Function where a path is computed such that latency variation in the path is minimized. In the context of loose hop expansion, the ERO expanding node MUST try to find a route such that overall latency variation of the loose hop is minimized.

2.2. Metric Bound subobject

The ERO subobject type Metric Bound (MB) is optional. It MAY be carried within an ERO object of RSVP-TE Path message and its scope is limited to previous ERO subobject that identifies an abstract node. It is possible to identify different Metric Bound subobjects for different hops of the ERO to be expanded. For more details please refer to the Processing Rules for the Metric Bound Subobjects section.

This subobject 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|L|    Type     |     Length    | metric-type |B|   Reserved    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          metric-bound                         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

The fields of the Metric subobject are defined as follows:

L bit: The L bit is set if the subobject represents a loose hop in the explicit route. If the bit is not set, the subobject represents a strict hop in the explicit route. Please note that use of MB subobject is also applicable to strict hops, e.g., in selecting a component link within a heterogeneous bundled TE link.

Type: The Type is to be assigned by IANA (suggested value: 67).

Length: The Length is 8.

Metric-type (8 bits): Specifies the metric type associated with the partial route expended by the node processing the loose ERO. The following values are currently defined:

* T=1: cumulative IGP cost

* T=2: cumulative TE cost

* T=3: Hop Counts

* T=4: Cumulative Latency

* T=5: Cumulative Latency Variation

B bit: Best-effort bit. When the best-effort (B) bit is set, it means that the ingress allows for the set up of an LSP that does not meeting the MB requirement. When the best-effort (B) bit is not set, it means that the MB needs to be observed.

Reserved: This field MUST be set to zero on transmission and MUST be ignored on receipt.

Metric-bound (32 bits): The metric-bound indicates an upper bound for the path metric that MUST NOT be exceeded for the ERO expending node to consider the computed path as acceptable. The metric bound is encoded in 32 bits using IEEE floating point format as defined in [IEEE.754.1985]). When it indicates a time value (i.e. Latency or Latency Variation) it is expressed in milliseconds.

2.3. Processing rules

A single OF subobjects SHOULD be used between a pair of abstract nodes in ERO. Multiple Metric Bound subobjects MAY be indicated for each hop to be expanded and MUST be placed after each abstract node subobject. Different Metric Bounds MAY be identified for each hop expansion.

2.3.1. Processing Rules for the OF Subobjects

The basic processing rules of an ERO are not altered. Please refer to [RFC3209] for details.

The scope of the OF subobject is the previous ERO subobject that identifies an abstract node, and the subsequent ERO subobject that identifies an abstract node. Multiple OF subobjects may be present between any pair of abstract nodes. However, only first OF subobject is analyzed and others are ignored.

The following conditions SHOULD result in Path Error with error code "Routing Problem" and error subcode "Bad EXPLICIT_ROUTE object":

. If the first OF subobject is not preceded by an ERO subobject identifying the next hop.
. If the OF subobject follows an ERO subobject identifying the next hop that does not have the L-bit set.

If the processing node does not understand the OF subobject, it SHOULD send a PathErr with the error code "Routing Error" and error value of "Bad Explicit Route Object" toward the sender [RFC3209].

If the processing node understands the OF subobject and the ERO passes the above mentioned sanity check and any other sanity checks associated with other ERO subobjects local to the node, the node takes the following actions:

. If the node supports the requested OF, the node expands the loose hop using the requested OF as optimization criterion for computing the route to the next abstract node. After processing, the OF subobjects are removed from the ERO. The rest of the steps for the loose ERO processing follow procedures outlined in [RFC3209].
. If the node understands the OF subobject but does not support the requested OF, it SHOULD send a Path Error with error code "Routing Problem" and a new error subcode "Unsupported Objective Function". The error subcode "Unsupported Objective Function" for Path Error code "Routing Problem" is to be assigned by IANA.
. If the OF is supported but policy does not permit applying it, the processing node SHOULD send a Path Error with error code "Policy control failure" (value 2) and subcode "objective function not allowed". The error subcode "objective function not allowed" for Path Error code "Policy control failure" is to be assigned by IANA.

2.3.2. Processing Rules for the MB subobject

The basic processing rules of an ERO are not altered. Please refer to [RFC3209] for details.

The scope of the MB subobject is between the previous ERO subobject that identifies an abstract node, and the subsequent ERO subobject that identifies an abstract node. Multiple MB subobjects may be present between any pair of abstract nodes.

If the processing node does not understand the MB subobject, it SHOULD send a PathErr with the error code "Routing Error" and error value of "Bad Explicit Route Object" toward the sender [RFC3209].

If the processing node understands the MB subobject and the ERO passes the above mentioned sanity check and any other sanity checks associated with other ERO subobjects local to the node, the node takes the following actions:

. For all the MB subobject(s), the node expands the ERO such that the requested metric bound(s) are met for the route between the two abstract nodes in the ERO. After processing, the Metric subobjects are removed from the ERO. The rest of the steps for the ERO processing follow procedure outlined in [RFC3209].
. If the node understands the MB subobject but cannot find a route to the next abstract node such that the requested metric bound(s) can be satisfied and the best-effort (B) bit is not set, it SHOULD send a Path Error with error code "Routing Problem" and a new error subcode "No route available toward destination with the requested metric bounds". The error subcode "No route available toward destination with the requested metric bounds" for Path Error code "Routing Problem" is to be assigned by IANA (See IANA section for details).
. If the node understands the Metric subobject but cannot find a route to the next abstract node such that the requested metric bound(s) can be satisfied and the best-effort (B) bit is set, it SHOULD send a Path Error message with error code "Notify Error" and a new error subcode "Route not matching the requested metric bounds" is to be assigned by IANA (See IANA section for details).
. The ERO expanding node SHOULD respect the Metric Bound constraints in realizing any segment recovery procedure to change the route of the segment expanded by the said node. If

best-effort (B) bit is set and the new recovery segment violates the Metric Bound constraints, the ERO expanding SHOULD send a Path Error message with error code "Notify Error" and a new error subcode "Route not matching the requested metric bounds" is to be assigned by IANA (See IANA section for details).

3. Security Considerations

This document does not introduce any additional security issues above those identified in [RFC5920], [RFC2205], [RFC3209], and [RFC3473].

4. IANA Considerations

4.1. ERO Subobject

This document adds the following two new subobject of the existing entry for ERO (20, EXPLICIT_ROUTE):

Value	Description
TBA (suggest value: 66)	Objective Function (OF) subobject
TBA (suggest value: 67)	Metric subobject

These subobject may be present in the Explicit Route Object, but not in the Route Record Object.

OF Code values carried in OF subobject requires an IANA entry with suggested values as defined in section 2.1.

4.2. New RSVP error sub-code

For Error Code = 24 "Routing Problem" (see [RFC2205]) the following sub-code is defined.

Sub-code	Value
No route available toward destination with the requested metric bounds	To be assigned by IANA. Suggested Value: TBA.

For Error Code = 25 "Notify Error" (see [RFC2205]) the following sub-code is defined.

Sub-code	Value
Route not matching the requested metric bounds	To be assigned by IANA. Suggested Value: TBA.

5. Acknowledgments

Authors would like to thank Matt Hartley, Ori Gerstel, Gabriele Maria Galimberti, Luyuan Fang and Walid Wakim for their review comments.

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.

[IEEE.754.1985] IEEE Standard 754, "Standard for Binary Floating-Point Arithmetic", August 1985.

[RFC4208] Swallow, G., Drake, J., Ishimatsu, H., and Y. Rekhter, "Generalized Multiprotocol Label Switching (GMPLS) User-Network Interface (UNI): Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Support for the Overlay Model", RFC 4208, October 2005.

6.2. Informative References

[RFC5920] Fang, L., Ed., "Security Framework for MPLS and GMPLS Networks", RFC 5920, July 2010.

[RFC5440] Vasseur, JP., Ed., and JL. Le Roux, Ed., "Path Computation Element (PCE) Communication Protocol (PCEP)", RFC 5440, March 2009.

[RFC5541] Le Roux, JL., Vasseur, JP., and Y. Lee, "Encoding of Objective Functions in the Path Computation Element Communication Protocol (PCEP)", RFC 5541, June 2009.

[ID-SERVICE-AWARE] D. Dhody, V. Manral, Z. Ali, G. Swallow, K. Kumaki, " Extensions to the Path Computation Element Communication Protocol (PCEP) to compute service aware Label Switched Path (LSP)", draft-ietf-pce-pcep-service-aware, work in progress.

[OSPF-TE-METRIC] S. Giacalone, D. Ward, J. Drake, A. Atlas, S. Previdi, "OSPF Traffic Engineering (TE) Metric Extensions", draft-ietf-ospf-te-metric-extensions, work in progress.

[ISIS-TE-METRIC] S. Previdi, S. Giacalone, D. Ward, J. Drake, A. Atlas, C. Filsfils, "IS-IS Traffic Engineering (TE) Metric Extensions", draft-previdi-isis-te-metric-extensions, work in progress.

Author's Addresses

Zafar Ali
Cisco Systems.
Email: zali@cisco.com

George Swallow
Cisco Systems.
swallow@cisco.com

Clarence Filsfils
Cisco Systems.
cfilsfil@cisco.com

Luyuan Fang
Cisco Systems.
lufang@cisco.com

Kenji Kumaki
KDDI Corporation
Email: ke-kumaki@kddi.com

Rudiger Kunze
Deutsche Telekom AG
Ruediger.Kunze@telekom.de

Daniele Ceccarelli
Ericsson
Email: daniele.ceccarelli@ericsson.com

Xian Zhang
Huawei Technologies
Email: zhang.xian@huawei.com