Extensions to Resource Reservation Protocol For Re-optimization of Loosely Routed Point-to-Multipoint Traffic Engineering LSPs

draft-ietf-mpls-p2mp-loose-path-reopt-01

Author list:

Tarek Saad (tsaad@cisco.com)

Rakesh Gandhi (rgandhi@cisco.com) - Presenter

Zafar Ali (zali@cisco.com)

Robert H. Venator (robert.h.venator.civ@mail.mil)

Yuji Kamite (y.kamite@ntt.com)

Outline

- **Scope and Requirements**
- **Problem Statement**
- **Signaling Extensions**
- **Update and Next Steps**

- **NOTE: This is delta from the last update in IETF-90 Toronto.**

Scope

- **P2MP-TE LSP [RFC4875]**
- **S2L Sub-LSP(s) signaled with Loose Hop ERO(s) or with no ERO [RFC3209]**
- **Loosely routed LSP re-optimization [RFC4736]**

Requirements

- As per P2MP-TE [RFC4875], an ingress node may:
  - Re-optimize the entire P2MP-TE LSP by resignaling all its S2L sub-LSP(s), i.e. all destinations.
  - Combine multiple Path/PathErr messages using S2L sub-LSP descriptor list to alleviate scale issue.
- A P2MP-TE LSP can use Path/PathErr messages defined in [RFC4736] for re-optimization of individual S2L sub-LSPs.

Agenda

- Scope and Requirements
- Problem Statement
- Signaling Extensions
- IETF Update and Next Steps

RFC4736 For P2MP-TE LSP Re-optimization Using Combined Messages - 1/2

- Combined Path message with a full list of S2L sub-LSPs in the descriptor list gets decomposed at branching LSRs.
- Only a subset of the S2L sub-LSPs added in the descriptor list of the Path message propagated to downstream mid-point LSRs.
- When a preferable path exists at such mid-point LSRs, the PathErr can only include the S2L sub-LSPs traversing that LSR.
- Issue: To infer which mode of re-optimization to invoke, i.e. sub-group based re-optimization using the same LSP-ID or tree based re-optimization using a different LSP-ID, additional logic is required.

  For example, waiting for some time to aggregate all possible PathErr messages before taking an action.
- Solution: Can be avoided by using the re-evaluation request messages for P2MP-TE LSP Tree re-optimization.

RFC4736 For P2MP-TE LSP Re-optimization Using Combined Messages - 2/2

- When a combined message may not be large enough to fit all S2L sub-LSPs, an LSR may fragment the large RSVP message.
- The ingress node may receive multiple PathErrs with sub-set of S2L sub-LSPs in each (either due to the combined Path message got fragmented or combined PathErr message got fragmented).
- The above leads to the same issue discussed on the last slide.
- Solution: Can be addressed by using markers to define a full set or subset of S2L sub-LSPs in the descriptor list.

Agenda

- Scope and Requirements
- Problem Statement
- Signaling Extensions
- IETF Update and Next Steps

Signaling Using Markers in Combined Messages

- When a Path message is not large enough to fit all S2L sub-LSPs in the descriptor list, an LSR may fragment the message.
- LSR MAY add optional S2L_SUB_LSP_MARKER_BEGIN and S2L_SUB_LSP_MARKER_END Objects at the beginning and at the end of the S2L sub-LSP descriptor list(s), respectively.
  - A mid-point LSR SHOULD wait to accumulate all S2L sub-LSPs before attempting to re-evaluate preferable path when a Path message for "Path Re-evaluation Request" is received with S2L_SUB_LSP_MARKER_BEGIN.
  - An ingress node SHOULD wait to accumulate all S2L sub-LSPs before attempting to trigger re-optimization when a PathErr message with "Preferable Path Exists" is received with S2L_SUB_LSP_MARKER_BEGIN.

Marker Objects in S2L_SUB_LSP Object

- **S2L_SUB_LSP_MARKER_BEGIN :**

  Class-Num 50, C-Type TBA by IANA

  +--------------------+--------------------+--+

  | Length (4 bytes)| Class_Num 50 | S2L_SUB_LSP_MARKER_BEGIN |

  +--------------------+--------------------+--+

- **S2L_SUB_LSP_MARKER_END :**

  Class-Num 50, C-Type TBA by IANA

  +--------------------+--------------------+--+

  | Length (4 bytes)| Class_Num 50 | S2L_SUB_LSP_MARKER_END |

  +--------------------+--------------------+--+

- **The S2L_SUB_LSP_MARKER_BEGIN Object is added before adding the first S2L_SUB_LSP_IPv4 or S2L_SUB_LSP_IPv6 Object in the S2L sub-LSP descriptor list.**
- **The S2L_SUB_LSP_MARKER_END Object is added after adding the last S2L_SUB_LSP_IPv4 or S2L_SUB_LSP_IPv6 Object in the S2L sub-LSP descriptor list.**

Agenda

- Scope and Requirements
- Problem Statement
- Signaling Extensions
- IETF Update and Next Steps

IETF Update and Next Steps

- **Document has been updated to address comments from the reviews as part of the WG adoption**
- **Welcome comments from the WG on the document especially on the changes presented today**
- **Request for early allocation for IANA code-points as draft has been (partly) implemented in our products**

Thank You.

Backup

Outline

- **Scope and Requirements**
- **Problem Statement**
- **Signaling Extension**

Scope

- **P2MP-TE LSP [RFC4875]**
- **S2L Sub-LSP(s) signaled with Loose Hop ERO(s) or with no ERO [RFC3209]**
- **Loosely routed LSP re-optimization [RFC4736]**

Requirements

As per P2MP-TE [RFC4875], an ingress node may:

1. **Re-optimize the entire P2MP-TE LSP by resignaling all its S2L sub-LSP(s), i.e. all destinations, OR,**
2. **Re-optimize individual S2L sub-LSP, i.e. individual destination.**

- **[RFC4875] does not define mechanisms to re-optimize loosely routed (inter-domain) P2MP-TE LSPs.**

Agenda

- Scope and Requirements
- Problem Statement
- Signaling Extension

RFC4736 P2P LSP Re-optimization

Addresses re-optimization of loosely routed P2P LSPs

1. **Ingress sends "Path Re-evaluation Request" to trigger evaluation at midpoint LSR expanding loose next hops.**
   - **flag (0x20) in SESSION_ATTRIBUTES object in the Path message.**
2. **The midpoint LSR sends a (un)solicited "Preferable Path Exists" to notify the ingress node to trigger re-optimization.**
   - **PathErr code 25 (notify error defined in [RFC3209]) with sub-code 6.**

- **[RFC4736] does not define mechanism for P2MP-TE LSP Re-optimization.**

(Re-using) RFC4736 for P2MP-TE LSP Re-optimization

- Ingress sends "Path Re-evaluation Request" (PRR) for each individual sub-LSP to trigger evaluation at midpoint LSR expanding loose next hops
  - Ingress may have to send path re-evaluation requests on all (100s) sub-LSP(s) to decide whether or not to re-optimize the whole P2MP-TE LSP
  - Ingress may have to "heuristically" wait and aggregate all responses for "better path exists" to decide whether or not to do per sub-LSP or per LSP re-optimization
    - Ingress may prematurely start per sub-LSP re-optimization and then decide to abort and perform LSP re-optimization
    - Ingress may prematurely start re-optimization of sub-set of sub-LSPs, that may result in data traffic duplication [RFC4875] [Section 14.2]
  - May produce undesired results when inter-operating due to timing related issues and different implementations
- Can be avoided by extending the re-evaluation request messages for P2MP-TE LSP Tree re-optimization.

(Re-using) RFC4736 for P2MP-TE LSP Re-optimization

- Midpoint LSR sends an (un)solicited "Preferable Path Exists" (PPE) for each individual sub-LSP to notify the ingress node to trigger re-optimization
  - Midpoint LSR can not differentiate whether the request is to evaluate per sub-LSP path or whole P2MP-TE tree
    - May have to "heuristically" accumulate received requests for all sub-LSPs (using a wait timer) to interpret this as a re-evaluation request for the whole P2MP-TE LSP Tree
    - May prematurely notify better path exists for a sub-set of S2L sub-LSPs
  - Midpoint LSR may have to send better path exists on all (100s) sub-LSP(s) when it determine a better P2MP-TE tree exists
  - May produce undesired results when inter-operating due to timing related issues and different implementations
- Can be avoided by extending the notify messages send by the midpoint LSR for P2MP-TE LSP Tree re-optimization.

Agenda

- Scope and Requirements
- Problem Statement
- Signaling Extension

Extensions For P2MP-TE LSP Tree Re-optimization

1. Ingress node sends “P2MP-TE Tree Re-evaluation Request" to query a a midpoint LSR for a preferable P2MP-TE LSP tree.
   - A new “P2MP-TE Tree Re-evaluation Request” flag is defined in Attributes Flags TLV of the LSP_ATTRIBUTES object [RFC5420] that is carried in a Path message
2. Midpoint LSR notifies ingress of solicited/unsolicited "Preferable P2MP-TE Tree Exists” node to trigger re-optimization of whole P2MP-TE LSP
   - Midpoint LSR sends a PathErr code 25 (notify error defined in [RFC3209]) with new sub-code "Preferable P2MP-TE Tree Exists”.
3. Any S2L sub-LSP of the LSP Tree transiting through the midpoint LSR can be selected to send the “P2MP-TE Tree Re-evaluation Request” to the midpoint LSR(s).
4. Notification of "Preferable P2MP-TE Tree Exists” can be sent back on the same S2L sub-LSP on which request was received on

Thank You.