Implementation Recommendations to Improve the Scalability of RSVP-TE Deployments

draft-beeram-teas-rsvp-te-scaling-rec-00
(*replaces* draft-beeram-mpls-rsvp-te-scaling-01)

Vishnu Pavan Beeram, Markus Jork (Juniper Networks)

Ina Minei (Google, Inc)

Rob Shakir (British Telecom)

Ebben Aries (Facebook)

Dante Pacella (Verizon)

Tarek Saad (Cisco Systems) Debut

Summary Of Changes

- Reworked the “Introduction” Section.
- Restructured the set of recommendations:
  - Pre-Requisites for implementing RI-RSVP and/or Per-Peer Flow Control.
  - Refresh Interval Independent RSVP (RI-RSVP).
  - Per-Peer Flow Control.
- Changed the “Intended Status” of the document:
  - Proposed Standard
- Added an Appendix for “Recommended Defaults”.

Document Focus

Makes a set of concrete implementation recommendations to improve the scalability of RSVP-TE deployments:

- Builds on all the scaling work and analysis done so far.
- Advocates the use of a couple of techniques to significantly cut down the amount of processing cycles required to maintain LSP state:
  - Refresh Interval Independent RSVP (RI-RSVP)
    - Eliminates RSVP's reliance on refreshes and refresh-timeouts.
  - Per-Peer Flow Control
    - Enables a busy RSVP speaker to apply back pressure to its peer(s).

RFC 2961 Specific Recommendations

- Base Pre-Requisites: An implementation supporting RI-RSVP and/or Per-Peer Flow Control:
  - SHOULD indicate support for RSVP RR extensions by default (override default via config).
  - MUST support reliable delivery of Path/Resv and the corresponding Tear/Err.
  - MUST support retransmit of all RSVP-TE messages using exponential backoff.
- Making Acknowledgements mandatory.
- Clarifications on reaching Rapid Retry Limit (Rl)
  - For Tear/Err, take no further action.
  - For Path/Resv,
    - SHOULD periodically retransmit until an Ack is received.
      - Recommended Periodic Retransmission Interval: 30 seconds.

Refresh Interval Independent RSVP (RI-RSVP)

An implementation that supports RI-RSVP:

- MUST support all the pre-requisites.
- MUST use a large value (10s of minutes) for the default configurable refresh interval (R).
  - Recommended Default Refresh Interval – 20 minutes.
- MUST implement coupling the state of individual LSPs with the state of the corresponding RSVP-TE signaling adjacency.
- MUST make use of Node-ID based Hello Session for detection of RSVP-TE signaling adjacency failures.
  - Recommended Default Node Hello Interval – 9 seconds.
- (If Bypass FRR is supported) MUST implement procedures specified in <draft-chandra-mpls-ri-rsvp-frr> to facilitate refresh-interval independent FRR.
- MUST indicate support for RI-RSVP via the CAPABILITY object in Hello messages.

Per-Peer Flow Control

An implementation that supports Per-Peer Flow Control:

- MUST support all the pre-requisites.
- MUST use lack of ACKs from a peer as an indication of peer's RSVP-TE control plane congestion.
- SHOULD use a Retry Limit (Rl) value of 7.
- SHOULD prioritize Tear/Err over trigger Path/Resv (messages that bring up new LSP state) sent to a peer when congestion is detected in the peer.
- MUST indicate support for Per-Peer Flow Control via the CAPABILITY object in Hello messages.

Other Recommendations

- If Bypass FRR is supported by an implementation, it SHOULD support the procedures discussed in <draft-mtaillon-mpls-summary-frr-rsvpte>.

Recommended Defaults

- Refresh-Interval (R) – **20** minutes
  - Refresh Timer randomly set in the range [10 m – 30 m]
- Node Hello-Interval – **9** seconds
  - Hello Timeout – 3.5*9 = 31.5 seconds
- Retry-Limit (Rl) – **7**
- Periodic Retransmission Interval (on reaching Rl) – **30** seconds

Next Steps

- Solicit feedback.
- Request WG Adoption.