

# RSVP-TE Scaling Recommendations

draft-beeram-mpls-rsvp-te-scaling

Vishnu Pavan Beeram, Yakov Rekhter (Juniper Networks)

Ina Minei (Google, Inc)

Ebben Aries (Facebook)

Rob Shakir (British Telecom)

Dante Pacella (Verizon)

# Introduction

- RSVP-TE [RFC3209] inherited and augmented standard RSVP [RFC2205].
  - As such, it inherited specific protocol characteristics that adversely affect control plane scalability:
    - Reliance on periodic refreshes for state synchronization.
    - Reliance on refresh timeout for stale state cleanup.
    - Lack of any flow-control mechanism.
- Further enhancements to RSVP (including Refresh Overhead Reduction techniques) and RSVP-TE have been made over the years.
- Draft discusses (via a cohesive set of scaling recommendations) how an implementation can use these enhancements to keep up with the increasing scaling demands.

# Recommendations: Eliminating reliance on refreshes

- Implement reliable delivery of Path/Resv messages [RFC2961].
- Indicate support for RSVP Refresh Overhead Reduction Extensions by default.
- Make the value of the refresh interval configurable with the default value of 20 mins.

# Recommendations: Eliminating reliance on refresh timeouts

In addition to the recommendations in the previous slide,

- Implement reliable delivery of Tear/Err messages [RFC2961].
- Implement coupling the state of individual LSPs with the state of the corresponding RSVP-TE signaling adjacency.
- Use Node-ID based Hello session [RFC4558] for detection of RSVP-TE signaling adjacency failures.
  - Use a default node hello-interval of 9 secs.
- Implement procedures specified in [draft-chandra-mpls-enhanced-frr-bypass] to facilitate refresh-interval independent FRR.

# Recommendations: Per Peer Flow Control

- Use lack of ACKs from a peer as an indication of peer's RSVP-TE control plane congestion and throttle messages to the affected peer.
- Retransmit of all RSVP-TE messages using exponential backoff.
- Increase the Retry Limit (RI) from 3 to 7.
- Prioritize Tear/Error over trigger Path/Resv sent to a peer when congestion is detected in the peer.

# Other Recommendations

- Making Acknowledgements mandatory.
- Clarifications on reaching Rapid Retry Limit (RI):
  - For Tear/Err, no further action.
  - For Path/Resv,
    - Start periodic retransmission (every 30 secs).
    - Periodic retransmission SHOULD continue until an ack is received.
- Avoiding use of Router Alert IP Option:
  - Use Bundle messages (addressed directly to neighbor).
- Checking Data Plane readiness:
  - Always check if the data plane of the new LSP is operational before moving traffic onto it.

# Next Steps

- Solicit feedback.