TEAS Working Group H. Sitaraman, Ed.
Internet-Draft V. Beeram
Intended status: Informational Juniper Networks
Expires: January 19, 2017 I. Minei
Google, Inc.
July 18, 2016

Recommendations for RSVP-TE and Segment Routing LSP co-existence
draft-sitaraman-sr-rsvp-coexistence-rec-00.txt

Abstract

Operators are looking to introduce services over Segment Routing (SR) LSPs in networks running Resource Reservation Protocol (RSVP-TE) LSPs. In some instances, operators are also migrating existing services from RSVP-TE to SR LSPs. For example, there might be certain services that are well suited for SR and need to co-exist with RSVP-TE in the same network. In other cases, services running on RSVP-TE might be migrated to run over SR. Such introduction or migration of traffic to SR might require co-existence with RSVP-TE in the same network for an extended period of time depending on the operator's intent. The following document provides solution options for keeping the traffic engineering database (TED) consistent across the network, accounting for the different bandwidth utilization between SR and RSVP-TE.

Copyright Notice

Table of Contents

1. Introduction

Introduction of SR [I-D.ietf-spring-segment-routing] in the same network domain as RSVP-TE [RFC3209] presents the problem of accounting for SR traffic and making RSVP-TE aware of the actual available bandwidth on the network links. RSVP-TE is not aware of how much bandwidth is being consumed by SR services on the network links and hence both at computation time (for a distributed computation) and at signaling time RSVP-TE LSPs will incorrectly place loads. This is true where RSVP-TE paths are distributed or centrally computed without a common entity managing both SR and RSVP-TE computation for the entire network domain.

The problem space can be generalized as a dark bandwidth problem to cases where any other service exists in the network that runs in parallel across common links and whose bandwidth is not reflected in the available and reserved values in the TED. While it is possible to configure RSVP-TE to only reserve up to a certain maximum link bandwidth and manage the remaining link bandwidth for other services, this is a deployment where SR and RSVP-TE are separated in the same network (ships in the night) and can lead to suboptimal link bandwidth utilization not allowing each to consume more, if required and constraining the respective deployments.

The high level requirements or assumptions to consider are:

1. Placement of SR LSPs in the same domain as RSVP-TE LSPs MUST not introduce inaccuracies in the TED used by distributed or centralized path computation engines.

2. Engines that compute RSVP-TE paths MAY have no knowledge of the existence of the SR paths in the same domain.

3. Engines that compute RSVP-TE paths SHOULD not require a software upgrade or change to their path computation logic.

4. Protocol extensions MUST be avoided or minimal as in many cases this co-existence of RSVP-TE and SR MAY be needed only during a transition phase.

5. Placement of SR LSPs in the same domain as RSVP-TE LSPs that are computed in a distributed fashion MUST not require migration to a central controller architecture for the RSVP-TE LSPs.

2. 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].

3. Solution options

3.1. Partitioning of static bandwidth

In this model, the static reservable bandwidth of an interface can be statically partitioned between SR and RSVP-TE and each can operate within that bandwidth allocation and SHOULD NOT preempt each other.

The downside of this approach is the inability to use the reservable bandwidth effectively and inability to use bandwidth left unused by the other protocol.

3.2. Centralized management of available capacity

In this model, a central controller performs path placement for both RSVP-TE signaled and SR LSPs. The controller manages and updates its own view of the in-use and the available capacity. As the controller is a single common entity managing the network it can have a unified and consistent view of the available capacity at all times.

A practical drawback of this model is that it requires the introduction of a central controller managing the RSVP-TE signaled LSPs as a prerequisite to the deployment of any SR-signaled LSPs. Therefore, this approach is not practical for networks where distributed TE with RSVP-TE signaled LSPs is already deployed, as it requires a redesign of the network and is not backwards compatible. This does not satisfy requirement 5.

Note that it is not enough for the controller to just maintain the unified view of the available capacity, it must also perform the path computation for the RSVP-TE LSPs, as the reservations for the SR LSPs are not reflected in the TED. This does not fit with assumption 2 mentioned earlier.

3.3. Flooding SR utilization in IGP

Using techniques in [RFC7810], [RFC7471] and [RFC7823], the SR utilization information can be flooded in IGP-TE and the RSVP-TE path computation engine (CSPF) can be changed to consider this information. This requires changes to the RSVP-TE path computation logic and would require upgrades in deployments where distributed computation is done across the network.

This does not fit with requirements 3 and 4 mentioned earlier.

3.4. Running SR over RSVP-TE

SR can run over dedicated RSVP-TE LSPs that carry only SR traffic. In this model, the LSPs can be one-hop or multi-hop and can provide bandwidth reservation for the SR traffic based on functionality such as auto-bandwidth. The model of deployment would be similar in nature to running LDP over RSVP-TE. This would allow the TED to stay consistent across the network and any other RSVP-TE LSPs will also be aware of the SR traffic reservations. In this approach, steps must be taken to prevent non-SR traffic from taking the SR-dedicated RSVP-TE LSPs, unless required by policy.

The drawback of this solution is that it requires SR to rely on RSVP-TE for deployment. Furthermore, the accounting accuracy/frequency of this method is dependent on performance of auto-bandwidth for RSVP-

TE. Note that for this method to work, the SR-dedicated RSVP-TE LSPs must be set up with the best setup and hold priorities in the network.

3.5. TED consistency by reducing RSVP-TE subscription

The solution relies on dynamically measuring SR traffic utilization on each TE interface and reducing the bandwidth allowed for use by RSVP-TE. It is assumed that SR traffic is higher priority than RSVP-TE and there can be different priority RSVP-TE LSPs. The following methodology can be used at every TE node for this solution:

- T: Traffic statistics collection interval

- N: Traffic averaging calculation (adjustment) interval such that N = k * T, where k is a constant multiplier

- RSVP-TE subscription percentage: The percentage of static bandwidth of an interface that is usable by RSVP-TE

- SR traffic threshold percentage: The percentage difference of traffic demand that when exceeded can result in a change to the RSVP-TE subscription percentage

- IGP-TE update threshold: Specifies the frequency at which IGP-TE updates should be triggered based on TE bandwidth updates on a link

At every interval T, each node SHOULD collect the SR traffic statistics for each of its TE interfaces. Further, at every interval N, given a configured SR traffic threshold percentage and a set of collected SR traffic statistics samples across the interval N, the SR traffic average (or any other traffic metric depending on the algorithm used) over this period is calculated.

If the difference between the calculated SR traffic average and the current SR traffic average (that was computed in the prior adjustment) is at least SR traffic threshold percentage, then the RSVP-TE subscription percentage for that interface MUST be adjusted. This MAY result in updates to maximum reservable link bandwidth in IGP-TE.

As SR traffic is considered higher priority compared to RSVP-TE, the reduction in RSVP-TE subscription percentage can result in RSVP-TE LSPs being hard or soft preempted. Such preemption will be based on relative priority (e.g. low to high) between RSVP-TE LSPs. It is RECOMMENDED that the IGP-TE update threshold SHOULD be lower in order to flood unreserved bandwidth updates often.

If LSP preemption is not acceptable, then the RSVP-TE subscription percentage cannot be reduced below what is currently reserved by RSVP-TE on that interface. This may result in bandwidth not being available for SR traffic. Thus, it is required that any external controller managing SR LSPs should be able to detect this situation (for example by subscribing to TED updates [RFC7752]) and should take action to reroute existing SR paths.

Generically, SR traffic (or any non-RSVP-TE traffic) should have its own priority allocated from the available priorities. This would allow SR to preempt other traffic according to the preemption priority order.

In this solution, the logic to retrieve the statistics, calculating averages and taking action to change the subscription percentages is an implementation choice, and all changes are local in nature.

The above solution offers the advantage of not introducing new network-wide mechanisms especially during scenarios of migrating to SR in an existing RSVP-TE network and reusing existing protocol mechanisms.

4. Acknowledgements

5. Contributors

The following individuals contributed to this document:

Chandra Ramachandran
Juniper Networks
Email: csekar@juniper.net

Raveendra Torvi
Juniper Networks
Email: rtorvi@juniper.net

Sudharsana Venkataraman
Juniper Networks
Email: sudharsana@juniper.net

6. IANA Considerations

This draft does not have any request for IANA.

7. Security Considerations

No new security issues are introduced in this document beyond is already part of RSVP-TE and Segment routing architectures.

8. References

8.1. Normative References

[I-D.ietf-spring-segment-routing]
Filsfils, C., Previdi, S., Decraene, B., Litkowski, S., and R. Shakir, "Segment Routing Architecture", draft-ietf-spring-segment-routing-09 (work in progress), July 2016.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <http://www.rfc-editor.org/info/rfc2119>.

[RFC3209] Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V., and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels", RFC 3209, DOI 10.17487/RFC3209, December 2001, <http://www.rfc-editor.org/info/rfc3209>.

8.2. Informative References

[RFC7471] Giacalone, S., Ward, D., Drake, J., Atlas, A., and S. Previdi, "OSPF Traffic Engineering (TE) Metric Extensions", RFC 7471, DOI 10.17487/RFC7471, March 2015, <http://www.rfc-editor.org/info/rfc7471>.

[RFC7752] Gredler, H., Ed., Medved, J., Previdi, S., Farrel, A., and S. Ray, "North-Bound Distribution of Link-State and Traffic Engineering (TE) Information Using BGP", RFC 7752, DOI 10.17487/RFC7752, March 2016, <http://www.rfc-editor.org/info/rfc7752>.

[RFC7810] Previdi, S., Ed., Giacalone, S., Ward, D., Drake, J., and Q. Wu, "IS-IS Traffic Engineering (TE) Metric Extensions", RFC 7810, DOI 10.17487/RFC7810, May 2016, <http://www.rfc-editor.org/info/rfc7810>.

[RFC7823] Atlas, A., Drake, J., Giacalone, S., and S. Previdi, "Performance-Based Path Selection for Explicitly Routed Label Switched Paths (LSPs) Using TE Metric Extensions", RFC 7823, DOI 10.17487/RFC7823, May 2016, <http://www.rfc-editor.org/info/rfc7823>.

Authors' Addresses

Harish Sitaraman (editor)
Juniper Networks
1133 Innovation Way
Sunnyvale, CA 94089
US

Email: hsitaraman@juniper.net

Vishnu Pavan Beeram
Juniper Networks
10 Technology Park Drive
Westford, MA 01886
US

Email: vbeeram@juniper.net

Ina Minei
Google, Inc.
1600 Amphitheatre Parkway
Mountain View, CA 94043
US

Email: inaminei@google.com