            Recommendations for RSVP-TE and Segment Routing (SR)
                  Label Switched Path (LSP) Coexistence

Abstract

   Operators are looking to introduce services over Segment Routing (SR)
   Label Switched Paths (LSPs) in networks running Resource Reservation
   Protocol - Traffic Engineering (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 coexist with RSVP-TE in the same network.
   Such introduction or migration of traffic to SR might require
   coexistence 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 consistent across the network, accounting for the different
   bandwidth utilization between SR and RSVP-TE.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

   Introduction of SR [RFC8402] 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; 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 Traffic Engineering Database
(TED). In most practical instances, given the static nature of the
traffic demands, limiting the reservable bandwidth available to RSVP-
TE has been an acceptable solution. However, in the case of SR
traffic, there is assumed to be very dynamic traffic demands, and
there is considerable risk associated with stranding capacity or
overbooking service traffic resulting in traffic drops.

The high-level requirements 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 should be avoided or be minimal as, in many
    cases, this coexistence 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", "NOT RECOMMENDED", "MAY", and
"OPTIONAL" in this document are to be interpreted as described in
BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all
capitals, as shown here.

3. Solution Options

The following section lists SR and RSVP coexistence solution options.
A specific solution is not recommended as all solutions are valid,
even though some may not satisfy all the requirements. If a solution
is acceptable for an operator based on their deployment model, then
such a solution can be chosen.

3.1. Static Partitioning of Bandwidth

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

   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 downside of this approach is the inability to use the reservable
   bandwidth effectively and the 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 and SR LSPs. The controller manages and updates its own view
   of the in-use and 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 LSPs as a
   prerequisite to the deployment of any SR LSPs. Therefore, this
   approach is not practical for networks where distributed TE with
   RSVP-TE 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.

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 (Constrained Shortest Path First (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, non-SR
   traffic MUST NOT take 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 Reflecting SR Traffic

   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 receives precedence in terms
   of the placement on the path over RSVP traffic (that is, RSVP traffic
   can be preempted from the path in case of insufficient resources).
   This is logically equivalent to SR traffic having the best preemption
   priority in the network. Note that this does not necessarily mean
   that SR traffic has higher QoS priority; in fact, SR and RSVP traffic
   may be in the same QoS class.

Reducing the bandwidth allowed for use by RSVP-TE can be explored using the three parameters available in IGP-TE ([RFC5305] [RFC3630]), namely Maximum-Link-Bandwidth, Maximum-Reservable-Bandwidth, and Unreserved-Bandwidth.

o Maximum-Link-Bandwidth: This parameter can be adjusted to accommodate the bandwidth required for SR traffic with cascading impacts on Maximum-Reservable-Bandwidth and Unreserved-Bandwidth. However, changing the maximum bandwidth for the TE link will prevent any compute engine for SR or RSVP from determining the real static bandwidth of the TE link. Further, when the Maximum-Reservable-Bandwidth is derived from the Maximum-Link-Bandwidth, its definition changes since Maximum-Link-Bandwidth will account for the SR traffic.

o Unreserved-Bandwidth: SR traffic could directly adjust the Unreserved-Bandwidth, without impacting Maximum-Link-Bandwidth or Maximum-Reservable-Bandwidth. This model is equivalent to the option described in Section 3.4. Furthermore this would result in overloading IGP-TE advertisements to directly reflect both RSVP-TE bandwidth bookings and SR bandwidth measurements.

o Maximum-Reservable-Bandwidth: As the preferred option, SR traffic could adjust the Maximum-Reservable-Bandwidth, with cascading impact on the Unreserved-Bandwidth.

The following methodology can be used at every TE node for this solution, using the following parameters:

o T: Traffic statistics collection time interval.

o k: The number of traffic statistics samples that can provide a smoothing function to the statistics collection. The value of k is a constant integer multiplier greater or equal to 1.

o N: Traffic averaging calculation (adjustment) interval such that N = k * T.

o Maximum-Reservable-Bandwidth: The maximum available bandwidth for RSVP-TE.

o If Diffserv-aware MPLS Traffic Engineering (DS-TE) [RFC4124] is enabled, the Maximum-Reservable-Bandwidth SHOULD be interpreted as the aggregate bandwidth constraint across all Class-Types independent of the Bandwidth Constraints model.

o Initial Maximum-Reservable-Bandwidth: The Maximum-reservable-
   bandwidth for TE when no SR traffic or RSVP-TE reservations exist
   on the interface.

o RSVP-unreserved-bandwidth-at-priority-X: Maximum-Reservable-
   Bandwidth - sum of (existing reservations at priority X and all
   priorities better than X).

o SR traffic threshold percentage: The percentage difference of
   traffic demand that, when exceeded, can result in a change to the
   RSVP-TE Maximum-Reservable-Bandwidth.

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

o M: An optional multiplier that can be applied to the SR traffic
   average. This multiplier provides the ability to grow or shrink
   the bandwidth used by SR. Appendix A offers further guidance on
   M.

At every interval T, each node SHOULD collect the SR traffic
statistics for each of its TE interfaces. The measured SR traffic
includes all labeled SR traffic and any traffic entering the SR
network over that TE interface. 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. This method of sampling traffic
statistics and adjusting bandwidth reservation accordingly is similar
to how bandwidth gets adjusted for auto-bandwidth RSVP-TE LSPs.

If the difference between the new 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 two
values MUST be updated:

o New Maximum-Reservable-Bandwidth = Initial Maximum-Reservable-
   Bandwidth - (new SR traffic average * M)

o New RSVP-unreserved-bandwidth-at-priority-X = New Maximum-
   Reservable-Bandwidth - sum of (existing reservations at priority X
   and all priorities better than X)

A DS-TE LSR that advertises a Bandwidth Constraints TLV should update
the bandwidth constraints for class-types based on operator policy.
For example, when Russian Dolls Model (RDM) [RFC4127] is in use, then

only BC0 may be updated. Whereas, when Maximum Allocation Model
(MAM) [RFC4125] is in use, then all Bandwidth Constraints (BCs) may
be updated equally such that the total value updated is equal to the
newly calculated SR traffic average.

Note that the computation of the new RSVP-unreserved-bandwidth-at-
priority-X MAY 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. The IGP-TE update threshold SHOULD allow
for more frequent flooding of unreserved bandwidth. From an
operational point of view, an implementation SHOULD be able to expose
both the configured and the actual values of the Maximum-Reservable-
Bandwidth.

If LSP preemption is not acceptable, then the RSVP-TE Maximum-
Reservable-Bandwidth 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 Maximum-Reservable-Bandwidth
is an implementation choice, and all changes are local in nature.
However, note that this is a new network trigger for RSVP-TE
preemption and thus is a consideration for the operator.

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. IANA Considerations

   This document has no IANA actions.

5. Security Considerations

   This document describes solution options for the coexistence of RSVP-
   TE and SR LSPs in the same administrative domain. The security
   considerations for SR are described in [RFC8402]. The security
   considerations pertaining to RSVP-TE are described in [RFC5920]. The

security considerations of each architecture are typically unaffected
by the presence of the other. However, when RSVP-TE and SR LSPs
coexist, it is possible for a hijacked SR traffic stream to
maliciously consume sufficient bandwidth and cause disruption to
RSVP-TE LSPs. With the solution option specified in Section 3.5, the
impact to RSVP-TE traffic can be controlled and paths re-routed.
Some latent risk of disruption still remains because this solution
option relies on taking statistics samples and adopting to new
traffic flows only after the adjustment period. The defensive
mechanisms described in the base SR security framework should be
employed to guard against situations that result in SR traffic
hijacking or denial of service.

6. References

6.1. Normative References

   [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate
              Requirement Levels", BCP 14, RFC 2119,
              DOI 10.17487/RFC2119, March 1997,
              <https://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,
              <https://www.rfc-editor.org/info/rfc3209>.

   [RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC
              2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174,
              May 2017, <https://www.rfc-editor.org/info/rfc8174>.

   [RFC8402] Filsfils, C., Ed., Previdi, S., Ed., Ginsberg, L.,
              Decraene, B., Litkowski, S., and R. Shakir, "Segment
              Routing Architecture", RFC 8402, DOI 10.17487/RFC8402,
              July 2018, <https://www.rfc-editor.org/info/rfc8402>.

6.2. Informative References

   [RFC3630] Katz, D., Kompella, K., and D. Yeung, "Traffic Engineering
              (TE) Extensions to OSPF Version 2", RFC 3630,
              DOI 10.17487/RFC3630, September 2003,
              <https://www.rfc-editor.org/info/rfc3630>.

   [RFC4124] Le Faucheur, F., Ed., "Protocol Extensions for Support of
              Diffserv-aware MPLS Traffic Engineering", RFC 4124,
              DOI 10.17487/RFC4124, June 2005,
              <https://www.rfc-editor.org/info/rfc4124>.

   [RFC4125] Le Faucheur, F. and W. Lai, "Maximum Allocation Bandwidth
              Constraints Model for Diffserv-aware MPLS Traffic
              Engineering", RFC 4125, DOI 10.17487/RFC4125, June 2005,
              <https://www.rfc-editor.org/info/rfc4125>.

   [RFC4127] Le Faucheur, F., Ed., "Russian Dolls Bandwidth Constraints
              Model for Diffserv-aware MPLS Traffic Engineering",
              RFC 4127, DOI 10.17487/RFC4127, June 2005,
              <https://www.rfc-editor.org/info/rfc4127>.

   [RFC5305] Li, T. and H. Smit, "IS-IS Extensions for Traffic
              Engineering", RFC 5305, DOI 10.17487/RFC5305, October
              2008, <https://www.rfc-editor.org/info/rfc5305>.

   [RFC5920] Fang, L., Ed., "Security Framework for MPLS and GMPLS
              Networks", RFC 5920, DOI 10.17487/RFC5920, July 2010,
              <https://www.rfc-editor.org/info/rfc5920>.

   [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,
              <https://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,
              <https://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,
              <https://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,
              <https://www.rfc-editor.org/info/rfc7823>.

Appendix A. Multiplier Value Range

   The following is a suggestion for the range of values for M:

   M is a per-node positive real number that ranges from 0 to 2 with a
   default of 1 and may be expressed as a percentage.

   o If M < 1, then the SR traffic average is being understated, which
      can result in the link getting full even though Maximum-
      Reservable-Bandwidth does not reach zero.

   o If M > 1, then the SR traffic average is overstated, thereby
      resulting in the Maximum-Reservable-Bandwidth reaching zero before
      the link gets full. If the reduction of Maximum-Reservable-
      Bandwidth becomes a negative value, then a value of zero SHOULD be
      used and advertised.

Acknowledgements

   The authors would like to thank Steve Ulrich for his detailed review
   and comments.

Contributors

   Chandra Ramachandran
   Juniper Networks
   Email: csekar@juniper.net

   Raveendra Torvi
   Juniper Networks
   Email: rtorvi@juniper.net

   Sudharsana Venkataraman
   Juniper Networks
   Email: sudharsana@juniper.net

   Martin Vigoureux
   Nokia
   Email: martin.vigoureux@nokia.com

Authors' Addresses

   Harish Sitaraman (editor)
   Juniper Networks
   1133 Innovation Way
   Sunnyvale, CA 94089
   United States of America

   Email: hsitaraman@juniper.net

   Vishnu Pavan Beeram
   Juniper Networks
   10 Technology Park Drive
   Westford, MA 01886
   United States of America

   Email: vbeeram@juniper.net

   Ina Minei
   Google, Inc.
   1600 Amphitheatre Parkway
   Mountain View, CA 94043
   United States of America

   Email: inaminei@google.com

   Siva Sivabalan
   Cisco Systems, Inc.
   2000 Innovation Drive
   Kanata, Ontario K2K 3E8
   Canada

   Email: msiva@cisco.com