SPRING Working Group
Internet-Draft
Intended status: Standards Track
Expires: December 9, 2018

C. Filsfils
Z. Ali, Ed.
Cisco Systems, Inc.
M. Horneffer
Deutsche Telekom
D. Voyer
Bell Canada
M. Durrani
Equinix
R. Raszuk
Bloomberg LP

June 7, 2018

Segment Routing Traffic Accounting Counters
draft-filsfils-spring-sr-traffic-counters-00.txt

Abstract

Segment Routing (SR) allows a headend node to steer a packet flow along any path. Intermediate per-flow states are eliminated thanks to source routing. Traffic accounting plays a critical role in network operation. A traffic account solution is required for SR networks that provides the necessary functionality without creating any additional per SR path states in the fabric.

This document describes counters for traffic accounting in SR networks.

Requirements Language

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 [RFC2119].

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at https://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any

time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on December 9, 2018.

Table of Contents

1. Introduction

This document defines counters for traffic accounting in segment routing (SR) [I-D.ietf-spring-segment-routing] networks. The essence of Segment Routing consists in scaling the network by only maintaining per-flow state at the source or edge of the network. Specifically, only the headend of an SR policy [I-D.filsfils-spring-segment-routing-policy] maintains the related per-policy state. Egress and midpoints along the source route do not maintain any per-policy state. The traffic counters described in this section respects the architecture principles of SR, while given visibility to the service provider for network operation and capacity planning.

This document specifies prefix-SID, interface and SR Policy counters to be implemented at each SR router. Furthermore, it describes the traffic matrix (TM) counters for accounting at the TM border routers (details described in Section 2.4).The goal of the document is to specify these necessary counters for traffic accounting in an SR network. The actual usage of this information and leveraging for various use-cases is outside the scope of this document. [I-D.ali-spring-sr-traffic-accounting] describes some of the use-cases and application of these counters in an SR network.

This document assumes that the routers export the traffic counters defined in Section 2 to an external controller. The methods for collection of this information by the controller is beyond the scope of the document.

2. SR Traffic Counters

2.1. Traffic Counters Naming convention

The section uses the following naming convention when referring to the various counters. This is done in order to assign mnemonic names to SR counters.

- The term counter(s) in all of the definitions specified in this document refers either to the (packet, byte) counters or the byte counter.
- SR: any traffic whose FIB lookup is a segment (IGP prefix/Adj segments, BGP segments, any type of segments) or the matched FIB entry is steered on an SR Policy.
- INT in name indicates a counter is implemented at a per interface level.
- E in name refers to egress direction (with respect to the traffic flow).
- I in name refers to ingress direction (with respect to the traffic flow).
- TC in name indicates a counter is implemented on a Traffic Class (TC) basis.
- TM in name refers to a Traffic Matrix (TM) counter.
- PRO in name indicates that the counter is implemented on per protocol/adjacency type basis. Per PRO counters in this document can either be accounts for:
  - LAB (Labelled Traffic): the matched FIB entry is a segment, and the outgoing packet has at least one label (that label does not have to be a segment label, e.g., the label may be a VPN label).
  - V4 (IPv4 Traffic): the matched FIB entry is a segment which is PoP'ed. The outgoing packet is IPv4.

  * V6 (IPv6 Traffic): the matched FIB entry is a segment which is PoP'ed. The outgoing packet is IPv6.

- POL in name refers to a Policy counter.
- BSID in name indicates a policy counter for labelled traffic.
- SL in name indicates a policy counter is implemented at a Segment-List (SL) level.

Counter nomenclature is exemplified using the following example:

- SR.INT.E.PRO: Per-interface per-protocol aggregate egress SR traffic.
- POL.BSID: Per-SR Policy labelled steered aggregate traffic counter.

2.2. Per-Interface SR Counters

For each local interface, node N maintains the following per-interface SR counters. These counters include accounting due to push, pop or swap operations on SR traffic.

2.2.1. Per interface, per protocol aggregate egress SR traffic counters (SR.INT.E.PRO)

The following counters are included under this category.

- SR.INT.E.LAB: For each egress interface (INT.E), N MUST maintain counter(s) for the aggregate SR traffic forwarded over the (INT.E) interface as labelled traffic.
- SR.INT.E.V4: For each egress interface (INT.E), N MUST maintain counter(s) for the aggregate SR traffic forwarded over the (INT.E) interface as IPv4 traffic (due to the pop operation).
- SR.INT.E.V6: For each egress interface (INT.E), N MUST maintain counter(s) for the aggregate SR traffic forwarded over the (INT.E) interface as IPv6 traffic (due to the pop operation).

2.2.2. Per interface, per traffic-class, per protocol aggregate egress SR traffic counters (SR.INT.E.PRO.TC)

This counter provides per Traffic Class (TC) breakdown of SR.INT.E.PRO. The following counters are included under this category.

- SR.INT.E.LAB.TC: For each egress interface (INT.E) and a given Traffic Class (TC), N SHOULD maintain counter(s) for the aggregate SR traffic forwarded over the (INT.E) interface as labelled traffic.

- SR.INT.E.V4.TC: For each egress interface (INT.E) and a given Traffic Class (TC), N SHOULD maintain counter(s) for the aggregate SR traffic forwarded over the (INT.E) interface as IPv4 traffic (due to the pop operation).

- SR.INT.E.V6.TC: For each egress interface (INT.E) and a given Traffic Class (TC), N SHOULD maintain counter(s) for the aggregate SR traffic forwarded over the (INT.E) interface as IPv6 traffic (due to the pop operation).

2.2.3. Per interface aggregate ingress SR traffic counter (SR.INT.I)

The SR.INT.I counter is defined as follows:

For each ingress interface (INT.I), N SHOULD maintain counter(s) for the aggregate SR traffic received on I.

2.2.4. Per interface, per TC aggregate ingress SR traffic counter (SR.INT.I.TC)

This counter provides per Traffic Class (TC) breakdown of the SR.INT.I. It is defined as follow:

For each ingress interface (INT.I) and a given Traffic Class (TC), N MAY maintain counter(s) for the aggregate SR traffic (matching the traffic class TC criteria) received on I.

2.3. Prefix SID Counters

For a remote prefix SID S, node N maintains the following prefix SID counters. These counters include accounting due to push, pop or swap operations on the SR traffic.

2.3.1. Per-prefix SID egress traffic counter (PSID.E)

This counter is defined as follows:

For a remote prefix SID S, N MUST maintain counter(s) for aggregate traffic forwarded towards S.

2.3.2. Per-prefix SID per-TC egress traffic counter (PSID.E.TC)

This counter provides per Traffic Class (TC) breakdown of PSID.E. It is defined as follows:

For a given Traffic Class (TC) and a remote prefix SID S, N SHOULD maintain counter(s) for traffic forwarded towards S.

2.3.3. Per-prefix SID, per egress interface traffic counter (PSID.INT.E)

This counter is defined as follows:

For a given egress interface (INT.E) and a remote prefix SID S, N SHOULD maintain counter(s) for traffic forwarded towards S over the (INT.E) interface.

2.3.4. Per-prefix SID per TC per egress interface traffic counter (PSID.INT.E.TC)

This counter provides per Traffic Class (TC) breakdown of PSID.INT.E. It is defined as follows:

For a given Traffic Class (TC), an egress interface (INT.E) and a remote prefix SID S, N MAY maintain counter(s) for traffic forwarded towards S over the (INT.E) interface.

2.3.5. Per-prefix SID, per ingress interface traffic counter (PSID.INT.I)

This counter is defined as follows:

For a given ingress interface (INT.I) and a remote prefix SID S, N MAY maintain counter(s) for the traffic received on I and forwarded towards S.

2.3.6. Per-prefix SID, per TC, per ingress interface traffic counter (PSID.INT.I.TC)

This counter provides per Traffic Class (TC) breakdown of PSID.INT.I. It is defined as follows:

For a given Traffic Class (TC), ingress interface (INT.I), and a remote prefix SID S, N MAY maintain counter(s) for the traffic received on I and forwarded towards S.

2.4. Traffic Matrix Counters

A traffic matrix T(N, M) is the amount of traffic entering the network at node N and leaving the network at node M, where N and M are border nodes at an arbitrarily defined boundary in the network [Traffic-Matrices] is. The TM border defines the arbitrary boundary nodes of a contiguous portion of the network across which service providers wish to measure traffic flows. The traffic matrix (also called demand matrix) contains all the demands crossing the TM border. It has as many rows as ingress edge nodes and as many columns as egress edge nodes at the TM border. The demand D(N, M) is the cell of the matrix at row N and column M. In other words, a Traffic Matrix provides, for every ingress point N into the network and every egress point M out of the network, the volume of traffic T(N, M) from N to M over a given time interval. To measure the traffic matrix, nodes in an SR network designate its interfaces as either internal or external.

When Node N receives a packet destined to remote prefix SID M, N maintains the following counters. These counters include accounting due to push, pop or swap operations.

2.4.1. Per-Prefix SID Traffic Matrix counter (PSID.E.TM)

This counter is defined as follows:

For a given remote prefix SID M, N SHOULD maintain counter(s) for all the traffic received on any external interfaces and forwarded towards M.

2.4.2. Per-Prefix, Per TC SID Traffic Matrix counter (PSID.E.TM.TC)

This counter provides per Traffic Class (TC) breakdown of PSID.E.TM. It is defined as follows:

For a given Traffic Class (TC) and a remote prefix SID M, N SHOULD maintain counter(s) for all the traffic received on any external interfaces and forwarded towards M.

2.5. SR Policy Counters

Per policy counters are only maintained at the policy head-end node. For each SR policy [I-D.filsfils-spring-segment-routing-policy] , the head-end node maintains the following counters.

2.5.1. Per-SR Policy Aggregate traffic counter (POL)

This counter includes both labelled and unlabelled steered traffic. It is defined as:

For each SR policy (P), head-end node N MUST maintain counter(s) for the aggregate traffic steered onto P.

2.5.2. Per-SR Policy labelled steered aggregate traffic counter (POL.BSID)

This counter is defined as:

For each SR policy (P), head-end node N SHOULD maintain counter(s) for the aggregate labelled traffic steered onto P. Please note that labelled steered traffic refers to incoming packets with an active SID matching a local BSID of an SR policy at the head-end.

2.5.3. Per-SR Policy, per TC Aggregate traffic counter (POL.TC)

This counter provides per Traffic Class (TC) breakdown of POL. It is defined as follows:

For each SR policy (P) and a given Traffic Class (TC), head-end node N SHOULD maintain counter(s) for the aggregate traffic (matching the traffic class TC criteria) steered onto P.

2.5.4. Per-SR Policy, per TC labelled steered aggregate traffic counter (POL.BSID.TC)

This counter provides per Traffic Class (TC) breakdown of POL.BSID. It is defined as follows:

For each SR policy (P) and a given Traffic Class (TC), head-end node N MAY maintain counter(s) for the aggregate labelled traffic steered onto P.

2.5.5. Per-SR Policy, Per-Segment-List Aggregate traffic counter (POL.SL)

This counter is defined as:

For each SR policy (P) and a given Segment-List (SL), head-end node N SHOULD maintain counter(s) for the aggregate traffic steered onto the Segment-List (SL) of P.

2.5.6. Per-SR Policy, Per-Segment-List labelled steered aggregate traffic counter (POL.SL.BSID)

This counter is defined as:

For each SR policy (P) and a given Segment-List (SL), head-end node N MAY maintain counter(s) for the aggregate labelled traffic steered onto the Segment-List SL of P. Please note that labelled steered traffic refers to incoming packets with an active SID matching a local BSID of an SR policy at the head-end.

3. Security Considerations

This document does not define any new protocol extensions and does not impose any additional security challenges.

4. IANA Considerations

This document has no actions for IANA.

5. Acknowledgement

The authors like to thank Kris Michielsen for his valuable comments and suggestions.

6. Contributors

The following people have contributed to this document:

Ketan Talaulikar
Cisco Systems
Email: ketant@cisco.com

Siva Sivabalan
Cisco Systems
Email: msiva@cisco.com

Jose Liste
Cisco Systems
Email: jliste@cisco.com

Francois Clad
Cisco Systems
Email: fclad@cisco.com

Kamran Raza
Cisco Systems
Email: skraza@cisco.com

Shraddha Hegde
Juniper Networks
Email: shraddha@juniper.net

Gaurav Dawra
LinkedIn
Email: gdawra.ietf@gmail.com

Rick Morton
Bell Canada
Email: rick.morton@bell.ca

Dirk Steinberg
Steinberg Consulting
Email: dws@steinbergnet.net

Bruno Decraene
Orange Business Services
Email: bruno.decraene@orange.com

Stephane Litkowski
Orange Business Services
Email: stephane.litkowski@orange.com

Luay Jalil
Verizon
Email: luay.jalil@verizon.com

7. References

7.1. Normative References

[I-D.filsfils-spring-segment-routing-policy]
Filsfils, C., Sivabalan, S., Hegde, S., daniel.voyer@bell.ca, d., Lin, S., bogdanov@google.com, b., Krol, P., Horneffer, M., Steinberg, D., Decraene, B., Litkowski, S., Mattes, P., Ali, Z., Talaulikar, K., Liste, J., Clad, F., and K. Raza, "Segment Routing Policy Architecture", draft-filsfils-spring-segment-routing-policy-06 (work in progress), May 2018.

[I-D.ietf-spring-segment-routing]
Filsfils, C., Previdi, S., Ginsberg, L., Decraene, B., Litkowski, S., and R. Shakir, "Segment Routing Architecture", draft-ietf-spring-segment-routing-15 (work in progress), January 2018.

[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>.

7.2. Informative References

[I-D.ali-spring-sr-traffic-accounting]
Ali, Z., Filsfils, C., Talaulikar, K., Sivabalan, S., Liste, J., Horneffer, M., Raszuk, R., Litkowski, S., Dawra, G., daniel.voyer@bell.ca, d., and R. Morton, "Traffic Accounting in Segment Routing Networks", draft-ali-spring-sr-traffic-accounting-01 (work in progress), May 2018.

[Traffic-Matrices]
Schnitter, T-Systems, S. and M. Horneffer, T-Com, "Traffic Matrices for MPLS Networks with LDP Traffic Statistics, Proc. Networks2004, VDE-Verlag 2004", 2015.

Authors' Addresses

Clarence Filsfils
Cisco Systems, Inc.

Email: cfilsfil@cisco.com

Zafar Ali (editor)
Cisco Systems, Inc.

Email: zali@cisco.com

Martin Horneffer
Deutsche Telekom

Email: martin.horneffer@telekom.de

Daniel Voyer
Bell Canada
671 de la gauchetiere W
Montreal, Quebec H3B 2M8
Canada

Email: daniel.voyer@bell.ca

Muhammad Durrani
Equinix

Email: mdurrani@equinix.com

Robert Raszuk
Bloomberg LP

Email: robert@raszuk.net