SPRING Working Group   Z. Ali
Internet-Draft   C. Filsfils
Intended status: Informational   K. Talaulikar
Expires: August 8, 2020   Siva Sivabalan
                                              Cisco Systems, Inc.
                                                    M. Horneffer
                                                Deutsche Telekom
                                                       R. Raszuk
                                                    Bloomberg LP
                                                    S. Litkowski
                                        Orange Business Services
                                                        D. Voyer
                                                       R. Morton
                                                     Bell Canada
                                                        G. Dawra
                                                        LinkedIn
                                               February 9, 2020

Traffic Accounting in Segment Routing Networks
draft-ali-spring-sr-traffic-accounting-04.txt

Status of this Memo

Abstract

Capacity planning is the continuous art of forecasting traffic
load and failures to evolve the network topology, its capacity,
and its routing to meet a defined Service-Level Agreement (SLA).
This document takes a holistic view of network capacity planning
and identifies the role of traffic accounting in network
operation and capacity planning, without creating any additional
states in the SR fabric.

Table of Contents

1 Introduction

Capacity planning is the continuous art of forecasting traffic
load and failures to evolve the network topology, its capacity,
and its routing to meet a defined Service-Level Agreement (SLA).

This document takes a holistic view of traffic accounting and
its role in operation and capacity planning in Segment Routing
(SR) networks.

One of the main architecture principles of Segment Routing (SR)
is that it maintains per-flow states only at the ingress nodes

to the SR domain. The approach taken in this document respects
the architecture principles of SR, i.e., this draft does not
create any additional control and data plane states at the
ingress, transit or egress node for traffic accounting. Only the
ingress node of an SR policy maintains per-flow counters for
traffic accounting, which are also needed for other use-cases
like billing.

The Traffic Matrix (TM) is one of the main components of the
holistic approach to traffic accounting taken in this document.
A network's traffic matrix is the volume of aggregated traffic
flows that enter, traverse and leave an arbitrarily defined
boundary in the network over a given time interval. 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 TM border defined for traffic matrix
collection does not have to be at the edge of the network, e.g.,
it can also be placed at the aggregation layer. Knowledge of the
traffic matrix is essential to efficient and effective planning,
design, engineering, and operation of any IP or MPLS network.

[I-D.draft-ietf-spring-segment-routing-policy] defines the
traffic matrix counters for accounting at the router. This draft
describes how these counters simplify traffic matrix collection
process. [I-D.draft-ietf-spring-segment-routing-policy] also
specifies policy, prefix-SID and interface counters for
accounting in an SR network. This document along with the
traffic counters defined in [I-D.draft-filsfils-spring-segment-
routing-policy] constitute the holistic view of traffic
accounting in an SR network.

This document assumes that the routers export the traffic
counters defined in [I-D.draft-filsfils-spring-segment-routing-
policy] to an external controller. It is also assumed that the
controller also collects the following information in order to
get the visibility required for traffic accounting:

  - Network topology information indicates all the nodes and their inter-
     connecting links (e.g. via BGP-LS [RFC7752]).
  - SR Policies instantiated at various node and their BSID (e.g. using
     PCEP as in RFC8231 or BGP-LS as in draft-ietf-idr-te-lsp-distribution).
  - Aggregate traffic counters and statistics for links that include link
     utilization, per TC statistics, drop counters, etc.
  - IPFIX data and the flow accounting information derived from it from an
     IPFIX collector.

The methods for collection of this information by the controller
is beyond the scope of the document.

2 SR Traffic Counters

[I-D.draft-ietf-spring-segment-routing-policy] specifies SR
counters that form building blocks on accounting in SR networks.
Listing all counters in this document is not the goal of this
draft. Some of those counters are outlined below:

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

For a remote prefix SID M, this counter accounts for the
aggregate traffic forwarded towards M.

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

This counter provides per Traffic Class (TC) breakdown of
PSID.E.

- Per-SR Policy Aggregate traffic counter (POL)

This counter accounts for both labelled and unlabeled traffic
steered on an SR policy (P). This counter is only maintained by
the head-end node.

Traffic matrix counters are outlined in the traffic matrix
section.

3 SR Traffic Matrix (TM)

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. 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.

3.1 TM Border

The service provider needs to establish Traffic Matrix (TM)
border to collect traffic matrix. The TM border defines the
boundary nodes of a contiguous portion of the network across

which the service provider wishes to measure traffic flows. The
TM border divides the network into two parts:

-      Internal part: a contiguous part of the network that is
located within the TM border.
-      External part: anything outside of the TM border

The TM border cuts through nodes, resulting in two types of
interfaces: internal and external interfaces. Interfaces are
internal if they are located inside the TM border, they are
external if they are found outside the TM border.

How a node marks it interfaces as external or internal is an
implementation matter and beyond the scope of this document.

  3.1 Choosing TM Border

An operator can choose where the TM border is located.
Typically, this will be at the edge of the network, but it can
also be placed at the aggregation layer. Or an operator can use
multiple TM borders for each of their network domains, with each
TM border cutting through different nodes; different TM borders
cannot cut through the same nodes.

  3.2 Deriving Demand Matrix

The goal is to measure the volume of traffic that enters a TM
border node n through an external interface and leaves through
an external interface of another TM border node m. This traffic
volume yields the traffic matrix entry $T_{n,m}$. Measuring this for
every pair of TM border nodes (n,m) results in the complete
traffic matrix.

Service providers use various techniques to compute traffic
matrix, including a combination of collecting link utilization,
gathering IPFIX data, collect MPLS forwarding statistics, etc.
A service provider may also use traffic matrix counters defined
in [I-D.draft-ietf-spring-segment-routing-policy] for this
purpose. The usefulness and applicability of the Traffic Matrix
do not depend on the TM collection mechanism.

  3.1 Traffic Matrix Counters

Traffic Matrix counters are defined in [I-D.draft-filsfils-
spring-segment-routing-policy]. The TM counters are summarized
in the following for completeness.

When Node N receives a packet, N maintains the following counters.

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

For a given remote prefix SID M, this counter accounts for all the traffic received on any external interfaces and forwarded towards M.

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

This counter provides per Traffic Class (TC) breakdown of PSID.E.TM.

4 Internet Protocol Flow Information Export (IPFIX)

Internet Protocol Flow Information Export (IPFIX) [RFC 7011]-[RFC7015] is a standard of export for Internet Protocol flow information. IPFIX is extensively deployed and used by network management systems to facilitate services such as measurement, security, accounting and billing. IPFIX also plays a vital role in traffic accounting in SR network. For example, IPFIX can be used for traffic accounting on an SR policy, without requiring any change to the SR-MPLS or IPFIX protocols.

5 Segment Routing Traffic Accounting

The SR counters, IPFIX data, Traffic Matrix, network topology information, node, and link statistics, SR policies configuration and various SR counters described in [I-D.draft-ietf-spring-segment-routing-policy], etc. constitute components of SR traffic accounting. This section describes some potential use of this information, but other mechanisms also exists.

One of the possible uses is centered around the traffic matrix. An external controller collects the traffic counters, including the traffic matrix, defined in [I-D.draft-filsfils-spring-segment-routing-policy] from the routers. Using the Traffic Matrix TM(N, M), the controller knows the exact traffic is entering node N and leaving node M, where node N and M are edge node on an arbitrary TM border. The controller also collects network topology and SR policies configuration from the network.

Using this information, the controller runs local path
calculation algorithm to map these demands onto the individual
SR paths. This enables a controller to determine the path that
would be taken through the network (including ECMP paths) for
any prefix at any node. Specifically, the controller starts with
distributing the TM(N, M) equally among all ECMP from node N to
node M. By repeating the process for all entry and exit nodes in
the network, the controller predicts how the demands are
distributed among SR paths in the network. The equal
distribution of the traffic demand assumption is validated by
correlating the projected load with the link and node statistics
and other traffic counters described in [I-D.draft-filsfils-
spring-segment-routing-policy]. Specifically, the various SR
counters described in [I-D.draft-filsfils-spring-segment-
routing-policy] provide the view of each segment's ingress and
egress statistics at every node and link in the network, which
is further supplemented by SR Policies' statistics that are
available at all head-end nodes. The uses this information to
adjust the predicted load, accordingly. How such adjustments are
performed is beyond the scope of this document. The predicted
traffic mapping to the individual SR path may be used for serval
purposes. That includes simulating what-if scenarios, develop
contingency and maintenance plans, manage network latency and to
anticipate and prevent congestion, etc. For example, if there is
congestion on the link between two nodes, the controller can
identify the SR path causing the congestion and how to re-route
it to relieve it.

Another possible use is built around the IPFIX data. IPFIX can
be used for traffic account on an SR policy, without requiring
any change to the SR-MPLS or IPFIX protocols. It provides a more
granular visibility of network flows (including SR Policy flows)
at any point in the network that can be correlated. For example,
IPFIX may be enabled on the nodes and links at the traffic
matrix border nodes to analyze the flows entering and leaving a
specific network region. Additionally, it can be also enabled at
any node or a specific link within the network for analyzing
flows through it either on demand or continuously basis. IPFIX
can also be enabled on the head-end nodes and endpoints of SR
Policies in the network to analyze flows steered through various
policies. When traffic is steered on an SR policy, the steering
is based on a match of the fields of the incoming packet. A
controller can replicate the matching criteria to account for
the traffic received at the egress for the given SR policy. The
policy counters, other traffic counters defined in
[I-D.draft-ietf-spring-segment-routing-policy], and information of
packet loss over policy can further supplement the IPFIX based

accounting for measuring, accounting, and billing on per policy
basis. Since IPFIX sampling also includes the MPLS label stack
on the packet and the underlying payload, the traffic flows for
a specific SR policy can also be determined at any intermediate
link or node in the network, if necessary.

Link level statistics information, derived using the ingress and
egress counters (including the QoS counters on a per TC basis),
provides the view of link utilization including for a specific
class of service at any point. This helps detect congestion for
the link as a whole or for specific class of service.

In summary, a controller can use the holistic view of traffic
accounting provided in this document to predicted traffic
mapping to the individual SR paths. The aggregate demands on the
network and their paths can be determined and correlated with
link utilization to identify the flows causing congestion for
specific links. Further visibility into all the flows on a link
can be achieved using the SR counters and supplemented by IPIX
data.

6 Security Considerations

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

7 IANA Considerations

This document does not define any new protocol or any extension
to an existing protocol.

8 References

8.1 Normative References

[I.D-ietf-spring-srv6-network-programming]
     Filsfils, C., et al., "Segment Routing Policy for Traffic
     Engineering", draft-ietf-spring-segment-routing-policy
     (work in progress), February 2020.

8.2. Informative References
[RFC7011] Specification of the IP Flow Information Export (IPFIX) Protocol for
     the Exchange of Flow Information. B. Claise, Ed., B. Trammell, Ed.,
     P. Aitken. September 2013. (Format: TXT=170852 bytes) (Obsoletes
     RFC5101) (Also STD0077) (Status: INTERNET STANDARD) (DOI:
     10.17487/RFC7011)

[RFC7012] Information Model for IP Flow Information Export (IPFIX). B. Claise,

      Ed., B. Trammell, Ed.. September 2013. (Format: TXT=50237 bytes)
      (Obsoletes RFC5102) (Status: PROPOSED STANDARD) (DOI:
      10.17487/RFC7012)

[RFC7013] Guidelines for Authors and Reviewers of IP Flow Information Export
      (IPFIX) Information Elements. B. Trammell, B. Claise. September
      2013. (Format: TXT=76406 bytes) (Also BCP0184) (Status: BEST CURRENT
      PRACTICE) (DOI: 10.17487/RFC7013)

[RFC7014] Flow Selection Techniques. S. D'Antonio, T. Zseby, C. Henke, L.
      Peluso. September 2013. (Format: TXT=72581 bytes) (Status: PROPOSED
      STANDARD) (DOI: 10.17487/RFC7014)

[RFC7015] Flow Aggregation for the IP Flow Information Export (IPFIX)
      Protocol. B. Trammell, A. Wagner, B. Claise. September 2013.
      (Format: TXT=112055 bytes) (Status: PROPOSED STANDARD) (DOI:
      10.17487/RFC7015)

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

9 Acknowledgments

  10 Contributors

Francois Clad
Cisco Systems, Inc.
fclad@cisco.com

Faisal Iqbal
Cisco Systems, Inc.
faiqbal@cisco.com

Authors' Addresses

Clarence Filsfils
Cisco Systems, Inc.
Email: cfilsfil@cisco.com

Zafar Ali Cisco Systems, Inc.
Email: zali@cisco.com

Ketan Talaulikar
Cisco Systems, Inc.
Email: ketant@cisco.com

Siva Sivabalan
Cisco Systems, Inc.
Email: msiva@cisco.com

Martin Horneffer
Deutsche Telekom
Email: martin.horneffer@telekom.de

Robert Raszuk
Bloomberg LP
Email: robert@raszuk.net

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

Daniel Voyer
Bell Canada
Email: daniel.voyer@bell.ca

Rick Morton
Bell Canada
Email: rick.morton@bell.ca
 Gaurav Dawra
LinkedIn
Email: gdawra.ietf@gmail.com