Segment Routing
Technology and Use Cases

Alberto Donzelli
Solution Architect  donz@cisco.com
6 Nov 2018
INTRODUCTION
Why Segment Routing

Network Resiliency
TI-LFA and automated 50ms protection

E2E Application Control - SDN ready
Assign the packet path as it exits the application

Service Velocity
Faster service creation, easier troubleshooting

Simplified Operations
Easy to operate the network

Network Simplification
Eliminate LDP, RSVP and other protocols

Simplified Service Creation
Easily configure L2VPN and L3VPN services

Enhanced User Experience
Faster response time, automated service creation

Standards Based
No vendor lock-in
Segment Routing Leading Use Cases

**Disjoint Path**
Ability to offer new services that provides location/power redundancy for their customers.

**Efficient Traffic Engineering**
Advanced Traffic Engineering is possible in a simplified manner with application defined paths.

**Flexible Algorithm (Flex –Algo)**
TE-path from anywhere to anywhere computed by IGP. Define own custom algorithm i.e. latency

**Low Latency Path**
Ability to provide a deterministic low latency path for Latency sensitive services.

**Unified Forwarding Plane**
Single data plane for packet transport in the network for operational simplicity

**Network slicing/Service Chaining**
Architecture optimised for service chaining of network functions and scalable soft slicing capabilities
Segment Routing
Segment Routing – Overall principle

Path expressed in the packet

Data

Dynamic path

Explicit path

Paths options

Dynamic (SPF computation)

Explicit (expressed in the packet)

Control Plane

Routing protocols with extensions (IS-IS, OSPF, BGP)

SDN controller (BGP, PCEP, NETCONF/YANG)

Data Plane

MPLS (segment labels)

IPv6 (+ SR extension header)

End to end forwarding behavior defined in the packet
Segment Routing Traffic Engineering

Mission — Route the luggage to Berlin via Mexico and Madrid

1. A unique and global luggage tag is attached to the luggage with the list of stops to the final destination

2. At each stop, the luggage is simply routed to the next hop listed on the luggage tag

RESULT: Path can be controlled Simple and scalable
SCALING: No state and limited label utilization MPLS LFIB with Segment Routing

- LFIB populated by IGP (ISIS / OSPF/BGP)
- Forwarding table remains constant (Nodes + Adjacencies) regardless of number of paths
Segment Routing – packet structure + segments

MPLS packet Format

- MPLS header
  - Active Segment
    - Label 1
    - Label 2
    - Label 3
    - Label 4
  - MPLS payload
    - TCP, UDP, QUIC

SRv6 packet format

- IPv6 header
  - Source Address
    - Locator 1
      - Function 1
  - Segment Routing Header
    - Active Segment
      - Locator 1
        - Function 1
      - Locator 2
        - Function 2
      - Locator 3
        - Function 3
  - IPv6 payload
    - TCP, UDP, QUIC

IGP Prefix segment
- IGP prefix
- Global identifier
- Signaled by ISIS/OSPF

IGP Adjacency segment
- Forward on IGP adjacency
- Local identifier
- Signaled by ISIS/OSPF
Segment Routing - packet structure + segments

**MPLS packet Format**
- **MPLS header**
  - Active Segment
  - Label 1
  - Label 2
  - Label 3
  - Label 4
- **MPLS payload**
  - TCP, UDP, QUIC

**SRv6 packet format**
- **IPv6 header**
  - Source Address
  - Locator 1 Function 1
- **Segment Routing Header**
  - Active Segment
  - Locator 1 Function 1
  - Locator 2 Function 2
  - Locator 3 Function 3
- **IPv6 payload**
  - TCP, UDP, QUIC

**IGP Prefix segment**
- IGP prefix
- Global identifier
- Signaled by ISIS/OSPF

**IGP Adjacency segment**
- Forward on IGP adjacency
- Local identifier
- Signaled by ISIS/OSPF

© 2017 Cisco and/or its affiliates. All rights reserved. Cisco Confidential
Segment Routing – packet structure + segments

MPLS packet Format

MPLS header
- Active Segment
  - Label 1
  - Label 2
  - Label 3
  - Label 4

MPLS payload
- TCP, UDP, QUIC

EPE Prefix segment
- Egress Peering Engineering
- Local Identifier
- Signaled by BGP-LS
Multiple SID lists Unequal load balancing

- A policy can have multiple SID list
- Each SID list has his own weight for unequal load balancing
Interoperability

Use Cases
Rapid Protection: automatic TI LFA FRR

<table>
<thead>
<tr>
<th>Problem</th>
<th>Solution</th>
<th>Benefits</th>
</tr>
</thead>
</table>
| Incomplete coverage, topology dependent coverage of classical LFA | Automated Topology Independent with guaranteed sub-50ms per-prefix protection | Simple and Automated  
IGP computed / No midpoint backup state  
Optimal  
Backup path following post-convergence path  
Scalable  
Cisco’s TI-FLA algorithm – optimized for scalability  
Post-convergence path computation and SID-list encoding |
Low Latency Path
Real time Delay Measurement

Applicability Examples

<table>
<thead>
<tr>
<th>Extreme</th>
<th>Voice</th>
</tr>
</thead>
<tbody>
<tr>
<td>Real-time Communications</td>
<td>Communications</td>
</tr>
<tr>
<td>Tactile Internet</td>
<td>Fixed / mobile</td>
</tr>
</tbody>
</table>

Solution

Compute Low Latency path based on measured link delay with MPLS Performance Monitoring (PM)

Benefits

Standard Based
MPLS Performance Management to measure real time link delay, jitter and loss
detect optical path reroute, by measuring delay/jitter/loss variations in real-time
Meet, Maintain and Monitor SLAs at all times

MPLS Performance Monitoring (PM)

Exhaustive Telemetry
ISIS/OSPF update upon significant change
Trigger SRTE re-optimization
Probe every 3s

Find the best delay-optimized path to Node 4

MPLS-PM (delay/jitter/loss)
Operate with Advanced Monitoring

**Problem**

**Traffic blackholes**
Very difficult to detect, troubleshoot
Ex: CEF looks OK, IGP looks OK, etc. but half the data plane traffic is dropped

**Solution**

**Data plane Monitoring**; a blackhole detection mechanism
SR router auto-tests its dataplane exactly as experienced by customer transit traffic

Real packets are injected through each combination of neighbors and destinations.
All the ECMP possibilities are tested one by one

**Benefits**

Proactive detection to enable fast reaction time to failures

Scalability when comparing with end-to-end probing techniques

---

Interested? SR Operations and Monitoring on [segment-routing.net](http://segment-routing.net)
Simplified planning: SR Traffic Matrix

Applicability Examples

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Capacity Planning</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Traffic Patterns</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Network Optimization</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IP or Optical</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Solution

With Cisco Segment Routing, the traffic/demand matrix collection is automated

No extra off-box tooling required

Benefits

- Capacity planning
- Bandwidth Optimization
- IP/Optical Optimization
SRTE SR Policy
SRTE with Segment Routing

Shortest ECMP Path

SRTE ECMP path

Classical TE

RSVP TE computation
SID List = {24014, 24048, 24059, 24017, 24073}
SR Policy Identification

• An SR Policy is uniquely identified by a tuple (head-end, color, end-point)
  Head-end: where the SR Policy is instantiated (implemented)
  Color: a numerical value to differentiate multiple SRTE Policies between the same pair of nodes
  End-point: the destination of the SR Policy

• At a given head-end, an SR Policy is uniquely identified by a tuple (color, end-point)
SR Policy Identification

- For the same destination different colors (*) for different SLA
- Green = Low Latency
- Blue = High Bandwidth

(*) color extended community is specified in RFC 5512
Segment Routing - Automated Steering (AS) based on destination

Automated steering directs traffic based on colors

Traffic for 10.10.10.1 - NH 8 and 20.20.20.1 - NH 8

Steer traffic into SR Policy based on destination
Segment Routing - On Demand Next Hop

- On Demand Next Hop
- 1 receive GREEN route
- No delay optimized path to 8
- HE compute and setup
- Route is Automatically steered

Steer traffic into SR Policy based on destination
Signaling SR Policy
candidate path via BGP

IPv4 – SR Policy
NPRI
Color green
End-point 4.4.4.4
Distinguisher 1234
Preferemce 200
Binding SID 4001
Segment List
Weight: 1
<16001, 16002, 30204>
Segment List
Weight: 1
<16003, 16004>

FIB on Node12:

<table>
<thead>
<tr>
<th>In</th>
<th>Out</th>
<th>Fraction</th>
</tr>
</thead>
<tbody>
<tr>
<td>4001</td>
<td>&lt;16001, 16002, 30204&gt;</td>
<td>50%</td>
</tr>
<tr>
<td></td>
<td>&lt;16003, 16004&gt;</td>
<td>50%</td>
</tr>
</tbody>
</table>
Flex Algo
SR IGP Flex Algo

- Complements the SRTE solution by adding new Prefix-Segments with specific optimization objective and constraints
  - minimize igp-metric or delay or te-metric
  - avoid SRLG or affinity
  - Plane A/B
  - Only encrypted links
Multiple Prefix SIDs for the same destination
Different paths!!

- What is SR IGP Flexible Algorithm?
- An operator-defined custom IGP algorithm leveraging dedicated Prefix-SID
- Example – operator configure the normal IGP and associate Pref-SID 16004 to Lo0
- operator defines Flex-Algo 129 as “minimize delay metric” with Prefix SID 18004 associated to Lo0
- For each destination two different SID are installed in routers FIB
Example Use-Case – Dual Plane

- Grey nodes support Algo 0/128/129
- Green nodes support 0/128
- Red nodes support 0/129
- Algo 128: minimize IGP metric
- Algo 129: minimize IGP metric
- Nodes advertise a Prefix SID for each Algo they support
  - For example, for node N: 16000 + N
    > + 0 for Algo 0
    > + 800 for Algo 128
    > + 900 for Algo 129
Thanks