Segment Routing Work Book by Orhan Ergun LLC

Orhan Ergun LLC

About the Author

Orhan Ergun, CCIE/CCDE Trainer, Author, Network Design Advisor and Cisco Champion 2019. Orhan Ergun is award winning Computer Network Architect, CCDE Trainer and Author. Orhan has well known industry certificates CCIE #26567 and CCDE #20140017.

Orhan has more than 17 years of networking experience and has been working on many medium and large-scale network design and deployment projects for Enterprise and Service Provider networks. He has been providing consultancy services to African, Middle East and some Turkish Service Providers and Mobile Operators for many years. Orhan has been providing Cisco network design training such as CCDE, Pre-CCDE, Service Provider Design and many advanced technologies for many years, and created best CCDE Training Program to share his network design experience and knowledge with the networking community. Orhan is sharing his articles and thoughts on his blog www.orhanergun.net. All the training and consultancy services related information can be found from his website. Orhan has a Training and Consultancy company located in Istanbul, Turkey

Rasoul Mesghali is a Cisco Certified Internetwork Expert (CCIE) #34938 (Routing & Switching) with over 12 years in the networking industry. Rasoul loves technology and never stop keeping up with the latest trend in technology. His experience includes Training, Consulting and Planning and deployment of MPLS/SR (Segment routing) and data center networks. He knows Python programming language very well and he does code and programming on a regular basis.

Vahid Tavajjohi is a Network Engineer with more than seven years of experience. He designs and operates network projects in large service providers and data centers. Knowledge of both data center and service provider technologies, Virtualization, Cloud, NFV, SDN, ZTP, and Scripting are main focuses of his career. Also, consultancy of large companies and service providers is key point for his sight in networks. He is a researcher and he is looking for new and edge technologies. Vahid have teaching experience of network courses, like service provider and data center for technical staff of companies.

Contents

Day1

Day-1 Segment Routing Fundamentals

Segment Routing Fundamentals

Segment Routing (SR) is a flexible, scalable way of doing source routing. The source chooses a path and encodes it in the packet header as an ordered list of segments. Segments are identifier for any type of instruction.

Each segment is identified by the segment ID (SID) consisting of a flat unsigned 32-bit integer. Segment instruction can be:

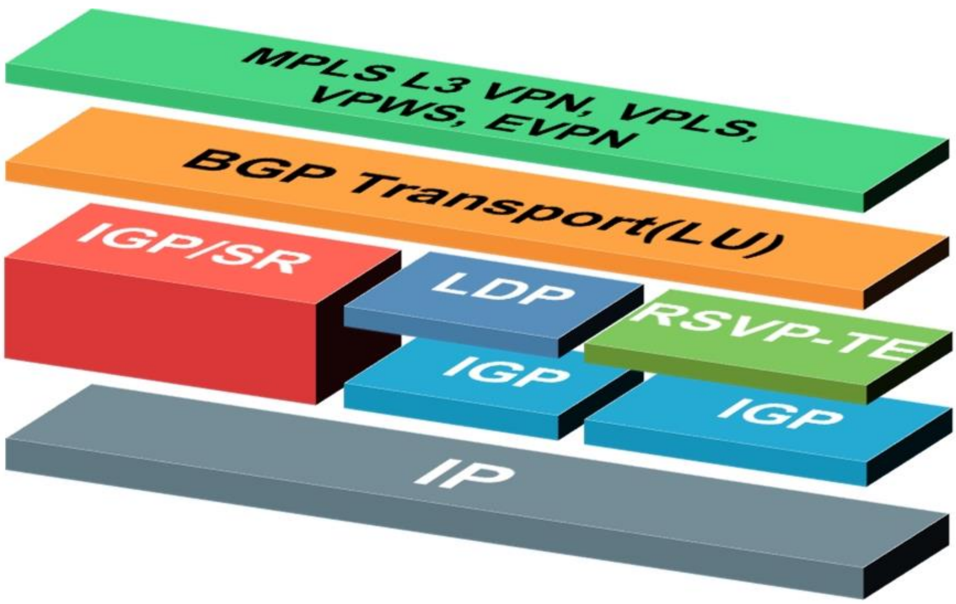
• Go to node N using the shortest path

• Go to node N over the shortest path to node M and then follow links Layer 1, Layer 2, and Layer 3

• Apply service S

With segment routing, the network no longer needs to maintain a per-application and per-flow state. Instead, it obeys the forwarding instructions provided in the packet.

As LDP and Segment-routing act as transport layer or underlay, VPLS, MPLS L3 VPN and EVPN are overlay services based on underlay technologies. In fact, any overlay technology can be carried by any underlay technology, for example, it is possible to provide l2VPN or EVPN services based on LDP or SR or RSVP-TE in transport layer.



This chapter is focused on SR fundamental concepts and differentiation of Underlay and overlay protocols. A Variety of LABs are provided based on ISIS and OSPF as IGP and traditional MPLS L3 VPN and VPLS as service layer. In the following scenarios you will learn how to configure sample SP network with SR in the core using ISIS/OSPF and provide MPLS L3 VPN and VPLS between CE routers.

The lab consist of combination of traditional IOS for CEs, IOS XE and IOS XR for SP routers. The following addressing table is applied on all labs in this chapter.

Addressing Table

Device Name	IPv4 Loopback address	Prefix-Sid
R1(XR)	1.1.1.1/32	16001
R2(XE)	2.2.2.2/32	16002
R3(XR)	3.3.3.3/32	16003
R4(XR)	4.4.4.4/32	16004
R5(XE)	5.5.5.5/32	16005
R6(XR)	6.6.6.6/32	16006
CE1(IOS)	111.111.111.111/32	NA
CE2(IOS)	222.222.222.222/32	NA

Addressing model between devices:

The address between nodes are using the following format:

10.10.XY.Z

X= Lower number, Y= Higher number, Z= Node number

For example:

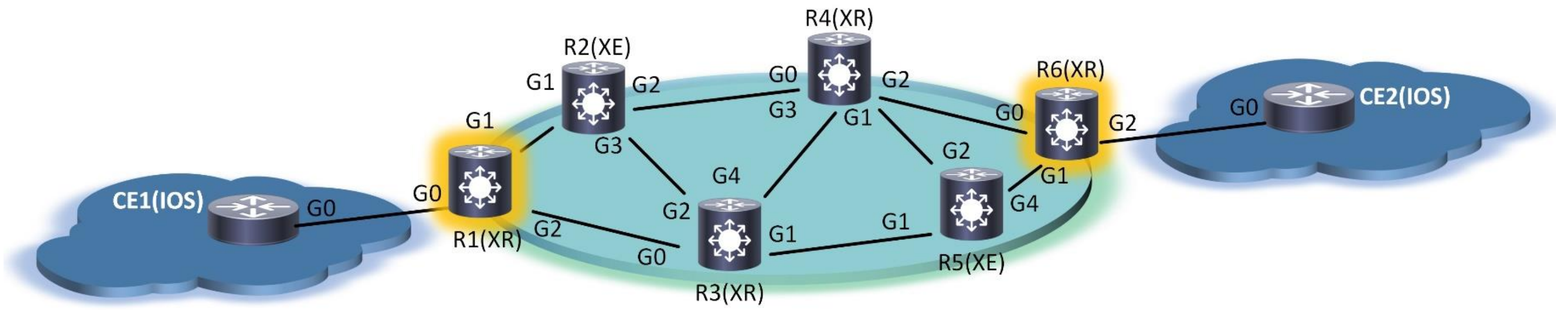
The address between R1, R2 =10.10.12.z

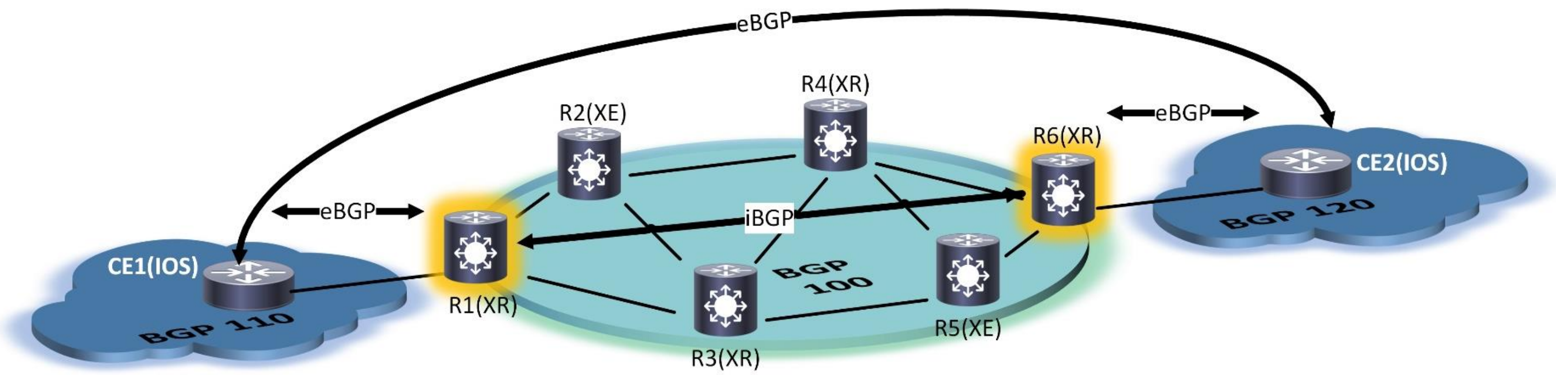
R1= 10.10.12.1, R2=10.10.12.2

Task1. Configure segment routing on all P and PE routers based on the following criteria:

a. configure ISIS as IGP (IPv4)
b. Level 2 only ISIS area
c. All routers are in 49.0000
d. Assign Prefix-SIDs based on SID table on loopback0
e. Configure all physical interface in the topology

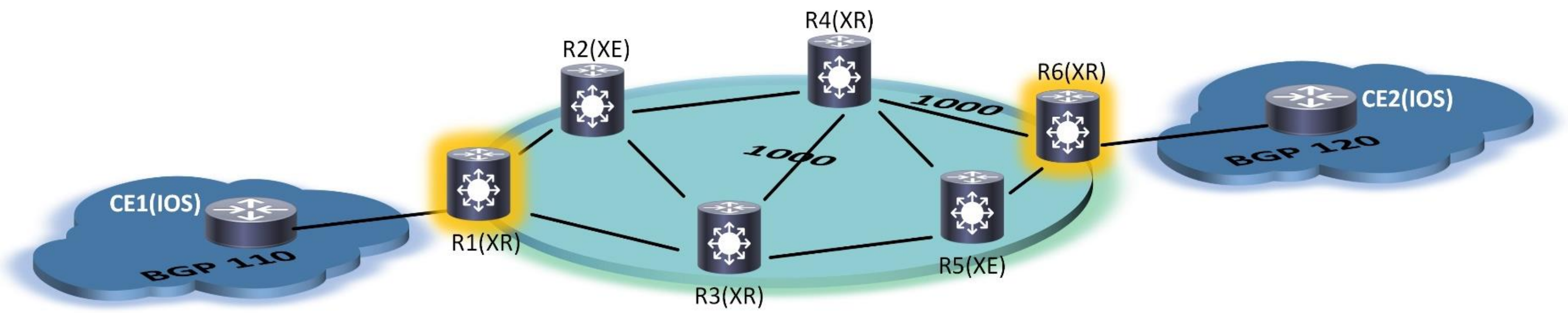
Topology





Task3.Configure LFA (IP-FRR) on all Routers in the map:

a. Configure IGP cost 1000 on the link between R3,R4 and R4,R6
b. Configure per-prefix LFA on all SP routers
c. Verify primary and backup path from R3 to R6 loopback



Configuration

LFA Configuration on R2 and R5 (XE):

```
router isis 1
fast-reroute per-prefix level-2 all
```

LFA Configuration on R1,R3,R4,R6 (XR)

```
Regarding topology, configure all physical interfaces as
follow:
router isis 1
interface GigabitEthernet0/0/0/X
  address-family ipv4 unicast
   fast-reroute per-prefix
```

Tip:

For directly connected per-prefix LFA, no additional label is imposed, the top label is swapped and packet is forwarded towards the LFA

Verification

On R2, verify protection coverage of per-prefix lfa using the comman **Show isis fast-reroute summary on R2**

```
RP/0/RP0/CPU0:R3#show isis fast-reroute summary
Mon Feb  4 07:52:15.649 UTC

IS-IS 1 IPv4 Unicast FRR summary

                          Critical   High       Medium     Low        Total
                          Priority   Priority   Priority   Priority
Prefixes reachable in L2
  All paths protected     0          0          5          0          5
  Some paths protected    0          0          0          0          0
  Unprotected             0          0          0          0          0
  Protection coverage     0.00%      0.00%      100.00%    0.00%      100.00%
```

```
RP/0/RP0/CPU0:R3#show isis fast-reroute 6.6.6.6/32
Mon Feb  4 07:53:15.654 UTC

L2 6.6.6.6/32 [20/115]
     via 10.10.35.5, GigabitEthernet0/0/0/1, R5, SRGB Base: 16000,
Weight: 0
       FRR backup via 10.10.34.4, GigabitEthernet0/0/0/3, R4, SRGB Base:
16000, Weight: 0, Metric: 1020
```

Verify the destinations that are protected via the low bandwidth, high cost link between R3 and R4 by using the command **show isis fast-reroute | include "L2|FRR backup via.*R4"**

Day2

Day-2 Segment Routing Internetworking with LDP

Segment Routing and LDP Coexistence

The Segment Routing architecture can be directly applied to the MPLS data plane with no change in the forwarding plane. In this chapter we are going to describe how Segment Routing operates in a network where LDP is deployed and in the case where SR-capable and non-SR-capable nodes coexist.

This chapter provides different scenarios in which you will learn a variety of the mechanisms through which SR interworks with LDP in cases where a mix of SR-capable and non-SR-capable routers co- exist within the same network and more precisely in the same routing domain. In addition, you will learn how to migrate your network from traditional LDP to Segment Routing.

There are some scenarios which show SR deployment can be used to provide SR benefits to LDP-based traffic including a possible application of SR in the context of inter-domain MPLS use-cases.

The lab consist of combination of traditional IOS for CEs, IOS XE and IOS XR for SP routers. The following addressing table is applied on all labs in this chapter.

Addressing Table

Device Name	IPv4 Loopback address	Prefix-Sid
R1(XR)	1.1.1.1/32	16001
R2(XE)	2.2.2.2/32	16002
R3(XR)	3.3.3.3/32	16003
R4(XR)	4.4.4.4/32	16004
R5(XE)	5.5.5.5/32	16005
R6(XR)	6.6.6.6/32	16006
CE1(IOS)	111.111.111.111/32	NA
CE2(IOS)	222.222.222.222/32	NA

Addressing model between devices:

The address between nodes are using the following format:

10.10.XY.Z

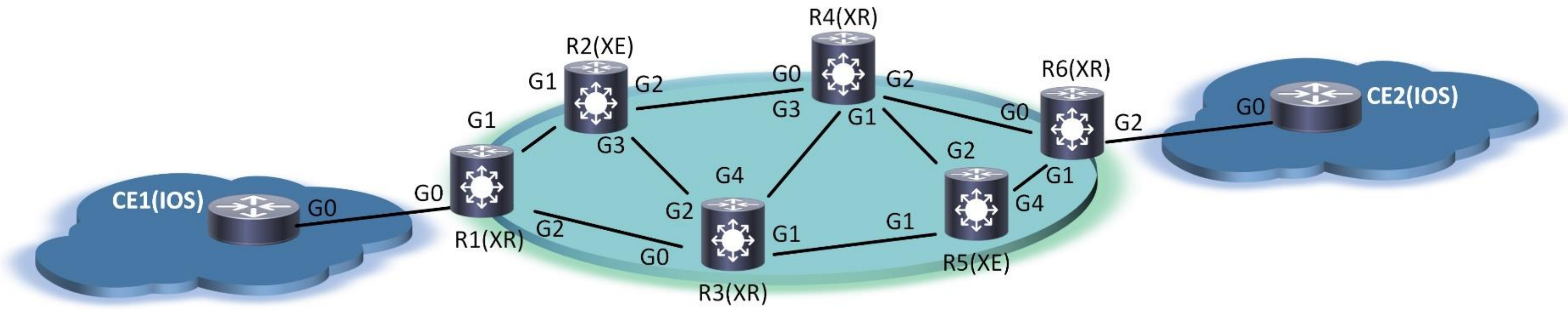
X= Lower number, Y= Higher number, Z= Node number

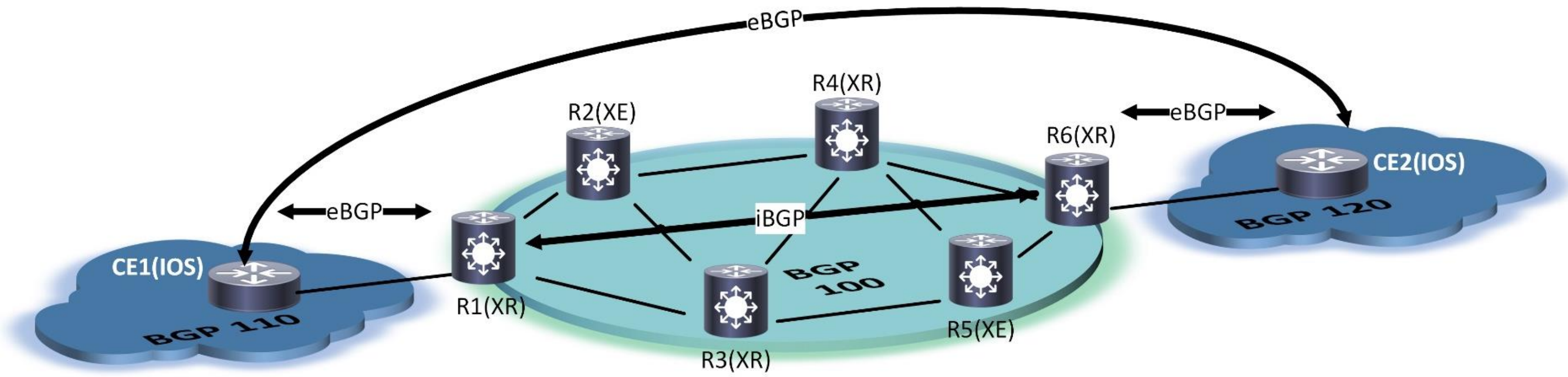
For example:

The address between R1, R2 =10.10.12.z

R1= 10.10.12.1, R2=10.10.12.2

Base Topology

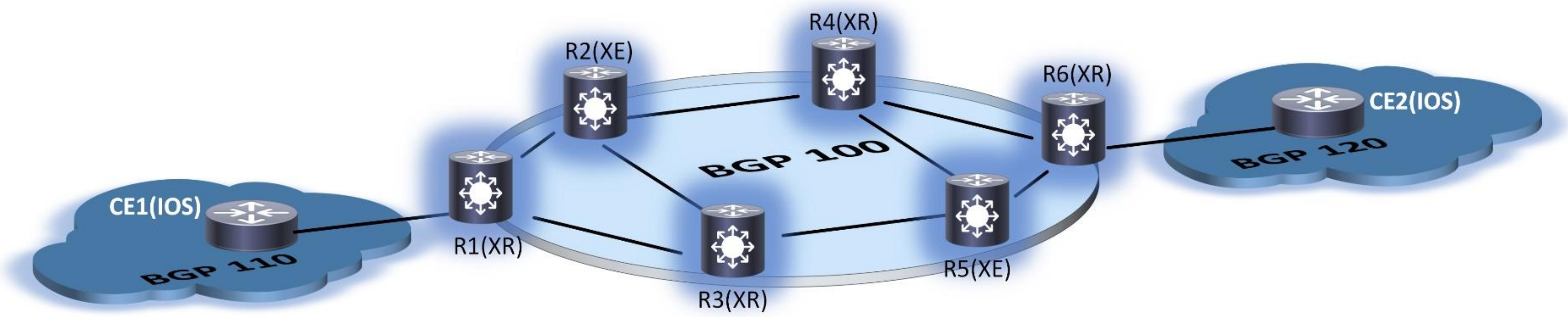




ICON	Capability
	SR + LDP Enabled Router
	LDP Only Router
	SR Only Router

Task1. Configure L3 VPN service and verify reachability between CEs:

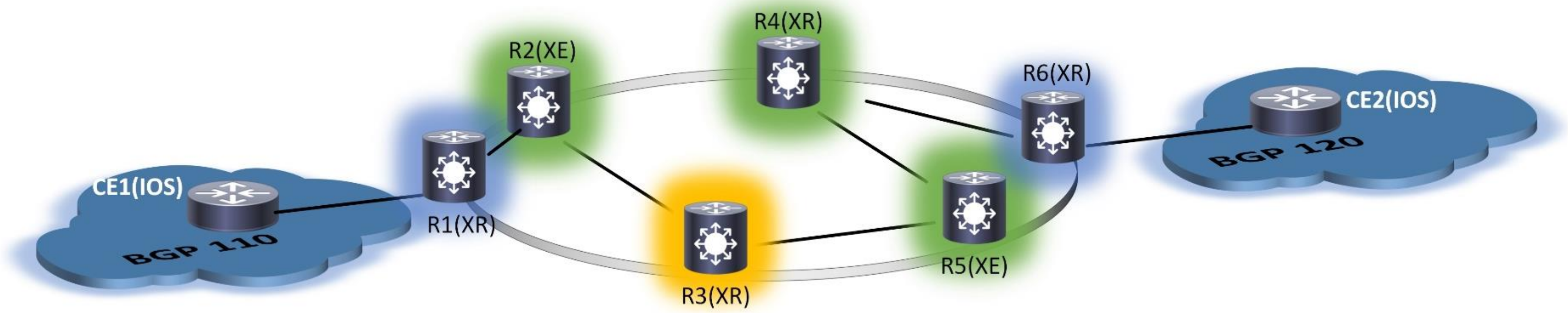
a. Configure ISIS as IGP and Configure LDP on all SP nodes
b. Put CE routers in VRF “A”
c. Site1 RT,RD: 100:1, Site2 RT,RD: 200:1
d. Assign AS 110 for site1 and AS120 for site2
e. Put PE routers in AS 100
f. Configure MP BGP on PEs
g. Advertise CE’s Loopbacks in MP-BGP
h. Configure BGP IPv4 session between CEs
i. CE1 and CE2 have a default route pointing to R1 and R6
j. Verify reachability CE’s loopback from remote CE using ping and traceroute



This step is to migrate to a complete segment routing topology (apart from some remaining LDP only nodes, R1, R6)

Task2. SR configuration as well as LDP

a) Configure segment routing on R2, R3, R4, R5 as well as LDP
b) Configure sr-prefer on R2 ,R3 ,R4, R5 to enable preference of Segment Routing over LDP for ip-to-mpls. Also, mapping server on node 3.
c) Changing topology to enforce segment routing inside core network.



Day3

Day-3 Segment Routing Traffic Engineering

Segment Routing Traffic engineering

SRTE has brought new Traffic Engineering techniques to tackle network operators scaling issues which prevented them in the past from having an end-to-end control over the variety of services they offer. In this chapter variety of scenarios are provided to cover SR-TE, automated steering, on demand next hop (ODN) and IGP SR Flexible Algorithm (Flex-Algo) in the single IGP domain.

Segment Routing steers a packet flow into SR Policy that contains an ordered list of segments. The Path can be different from the least cost path. Encode path information in the packet. A SR Policy is a framework that enables instantiation of an ordered SID list on a node for implementing a source routing policy and it is uniquely identified through a tuple (Head-end, color, and Endpoint). SR policy also can be used for Fast Reroute (FRR) or Operations, Administration, and Maintenance (OAM) purposes. Compared to RSVP-TE, advantages of SRTE are Multi-domain support by using PCEP for computation, ECMP/WECMP and Automated steering traffic. Also, there is a component named Binding-SID (B-SID) that involves a list of SIDs and it bound to SR Policy for greater scalability.

The head-end imposes the corresponding MPLS label stack on to outgoing packets to be carried over the tunnel. Each transit node along the SR-TE LSP path uses the incoming top label to select the next-hop, pop or swap the label, and forward the packet to the next node with the remainder of the label stack, until the packet reaches the ultimate destination. OSPF/ISIS provides TE with the topology and SR related information. SR related information include SRGB/prefix/Adjacency SIDs of all nodes/links with SR enabled in the network.

Correspondingly, the Automate Steering using BGP community is another technique which helps steering service traffic in a SR-policy that is covered in this chapter. ODN and Flex-Algo scenarios are provided to help you bringing scalability and flexibility to your network.

There are two types of SR-TE policies: dynamic and explicit.

Dynamic SR-TE Policy

When you configure local dynamic SR-TE, the head-end locally calculates the path to the destination address. Dynamic path calculation results in a list of interface IP addresses that traffic engineering (TE) maps to adj-SID labels. Routes are learned by way of forwarding adjacencies over the TE tunnel.

Explicit SR-TE Policy

An explicit path is a list of IP addresses or labels, each representing a node or link in the explicit path. This feature is enabled through the explicit-path command that allows you to create an explicit path and enter a configuration submode for specifying the path.

Note: It should be noted that both types explicit and dynamic policy can be applied using PCE controller in which external controller gathers topology information of multiple domains and select the best path over multiple domains and configures head end router to use injected segment list.

Addressing Table

Device Name	IPv4 Loopback address	Prefix-Sid
R1(XR)	1.1.1.1/32	16001
R2(XE)	2.2.2.2/32	16002
R3(XR)	3.3.3.3/32	16003
R4(XR)	4.4.4.4/32	16004
R5(XE)	5.5.5.5/32	16005
R6(XR)	6.6.6.6/32	16006
CE1(IOS)	111.111.111.111/32	NA
CE2(IOS)	222.222.222.222/32	NA

Addressing model between devices:

The address between nodes are using the following format:

10.10.XY.Z

X= Lower number, Y= Higher number, Z= Node number

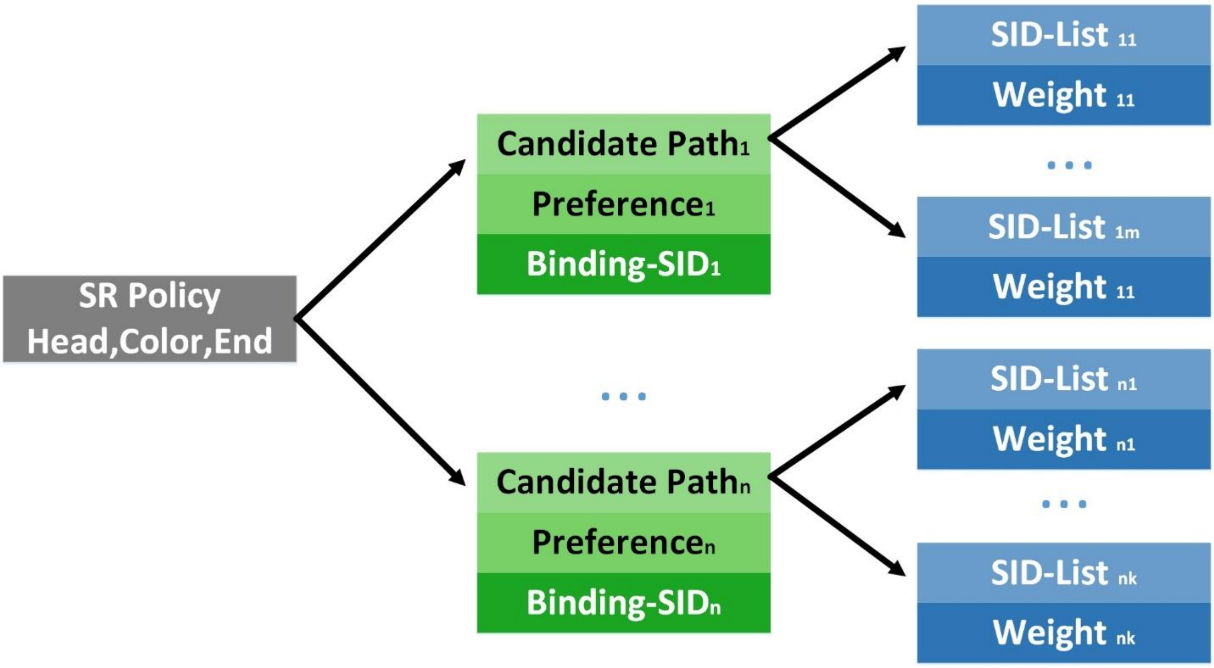
For example:

The address between R1, R2 =10.10.12.z

R1= 10.10.12.1, R2=10.10.12.2

Tip:

Candidate path has a preference and is associated with a single Binding-SID.

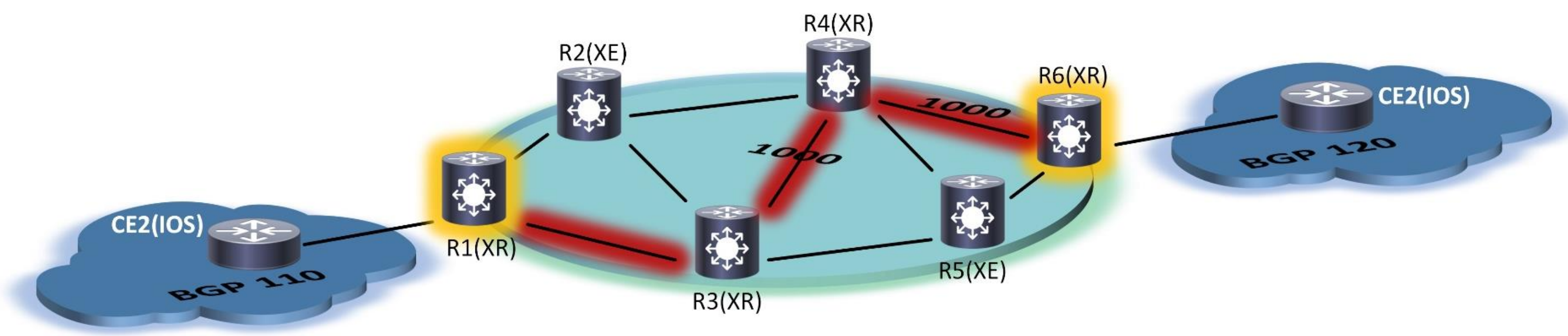


Tip:

SRTE Candidate Path can be received from different sources such as: CLI, BGP, PCEP, Netconf. Source of path is not considered for path selection. The valid path with Higher preference is the selected path.

Task1. Single domain static SRTE:

k. Configure ISIS1 as IGP and Configure SR on all SP nodes
l. Change the IGP metric of links <R3-R4> and <R4-R6> to 1000
m. On R3, configure SRLB to assign persistence Adj-SID and verify the Adj-SID and for R4, use dynamic Adj-SID
n. Configure R1 as Head-end and SRTE to use Explicit-Path through R3<->R4 link and assign BSID 40000 to the SR-Policy
o. R1 must reach R6 through the path R1-R3-R4-R6
p. On R1, configure TI-LFA
q. Check the TI-LFA backup path for the SRTE path on R1
r. Shutdown R3<->R4 link. verify that traffic will drop because of invalidation drop



a. Configure ISIS1 as IGP and Configure SR on all SP nodes

Day4

Day-4 Segment Routing Multi-Domain SRTE

Multi-Domain SRTE

You can scale your network with segment routing traffic engineering.

In segment routing it is possible like traditional unified MPLS (seamless MPLS) to use inter-AS options and BGP-LU for multi domain interconnection and in this chapter some labs are provided which shows how to use traditional methods using Segment Routing. In the meantime, an external controller (PCE Controller) using BGP-LS can gather topology information of different domains and provide the best end-to-end path using dynamic or explicit path policy. BGP-LS is an extension to Border Gateway Protocol (BGP) for distributing the network's link-state (LS) topology model to external entities, such as the SDN controller. It has received a lot of attention because many SDN apps need this model BGP-LS supports IS-IS and OSPFv2 (until the time of writing this document).

Segment routing for traffic engineering (SR-TE) uses a "policy" to steer traffic through the network. An SR-TE policy path is expressed as a list of segments that specifies the path, called a segment ID (SID) list. Each segment is an end-to-end path from the source to the destination, and instructs the routers in the network to follow the specified path instead of the shortest path calculated by the IGP. If a packet is steered into an SR-TE policy, the SID list is pushed on the packet by the head-end. The rest of the network executes the instructions embedded in the SID list.

In this chapter a number of useful multi domain scenarios are provided which can be usable for mobile operators and internet service providers.

Addressing Table

Device Name	IPv4 Loopback address	Prefix-Sid
R1(XR)	1.1.1.1/32	16001
R2(XR)	2.2.2.2/32	16002
R3(XR)	3.3.3.3/32	16003
R4(XR)	4.4.4.4/32	16004
R5(XR)	5.5.5.5/32	16005
R6(XR)	6.6.6.6/32	16006
RR1	7.7.7.7/32	
RR2	8.8.8.8/32	
CE1(IOS)	111.111.111.111/32	NA
CE2(IOS)	222.222.222.222/32	NA

Addressing model between devices:

The address between nodes are using the following format:

10.10.XY.Z

X= Lower number, Y= Higher number, Z= Node number

For example:

The address between R1, R2 =10.10.12.z

R1= 10.10.12.1, R2=10.10.12.2

Task1.Multidomain for SR and Dynamic SRTE using PCE

a) Configure two ISIS domains, ISIS1+SR and OSPF1+SR
b) Configure BGP AS100 in ISIS1 and AS200 in OSPF1
c) Configure ebgp between RR1 and RR2.
d) Configure RR1 and RR2 as XTC+PCE and R1, R6 PCC.
e) Configure mpls traffic-engineering under all nodes.
f) Put CE routers in VRF “A” and site1 RT,RD: 100:1, Site2 RT,RD: 200:1
g) Assign AS 110 for site1 and AS120 for site2 and advertise CE’s Loopbacks in MP-BGP