Segment routing (SR)

Definition

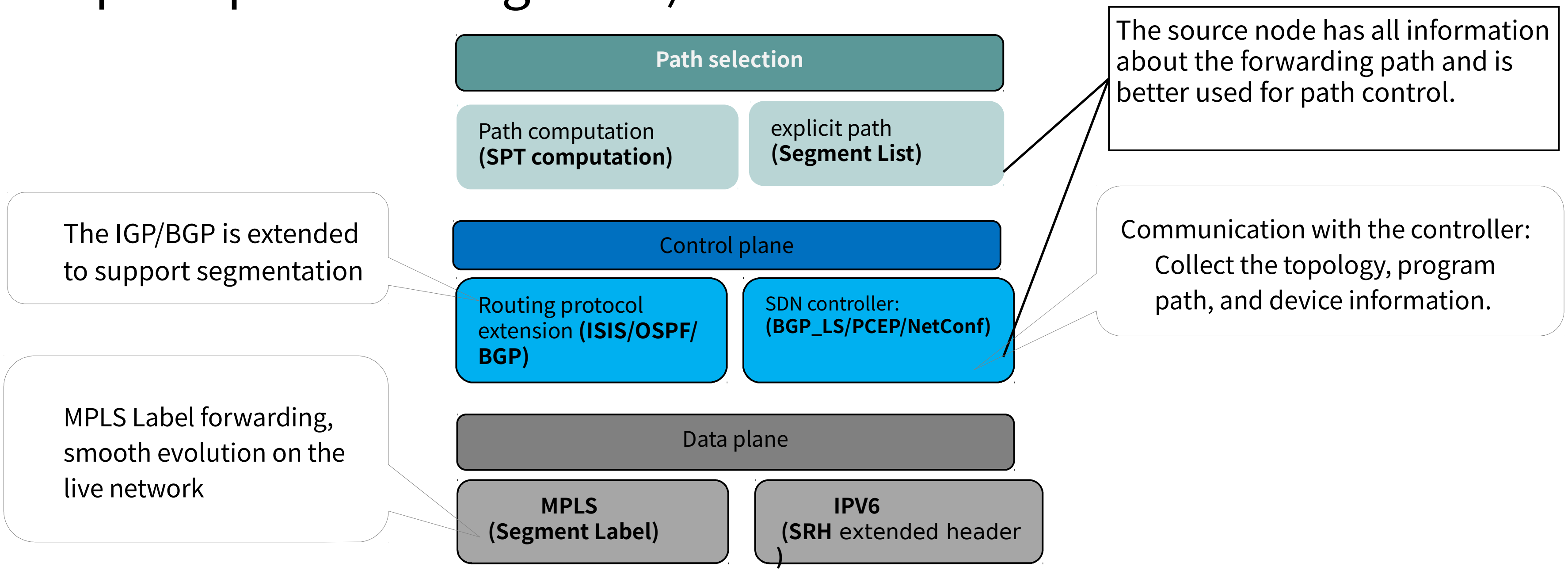
- Segment Routing (SR) is a protocol designed to forward data packets on a network based on source routes.
- Segment Routing divides a network path into several segments and assigns a segment ID (SID) to each segment and forwarding node. The segments and nodes are sequentially arranged into a segment list to form a forwarding path.

Segment Routing Fundamentals

- **Segment Routing domain**: a set of Segment Routing nodes.
- **SID**: unique identifier of a segment. A SID is mapped to an MPLS label on the forwarding plane.
- **Segment Routing global block (SRGB)**: a set of local labels reserved for Segment Routing.
- **Segment List**: It is a sort set of destination address prefix SID / node SID and adjacent SID ordered list, and is used to identify a complete label switching path LSP(Label Switched Path). In the MPLS architecture, the segment list is encapsulated in the packet header to guide forwarding.

SR Background -SR Technical Framework

- SR is a technical solution of packet forwarding provided by adding a series of segment identifiers to a packet only on the source (an explicit path loading node) node



Segment Routing Fundamentals

- **Prefix segment** (Manually configured) - Identifies the prefix of a destination address.
- An IGP propagates the prefix segment of an NE to the other
- NEs. The prefix segment is visible and takes effect globally. Each prefix segment is identified by a prefix SID. A prefix SID is an offset value within the SRGB range and advertised by a source node. The receive end uses the local SRGB to compute label values and then generates MPLS forwarding entries.

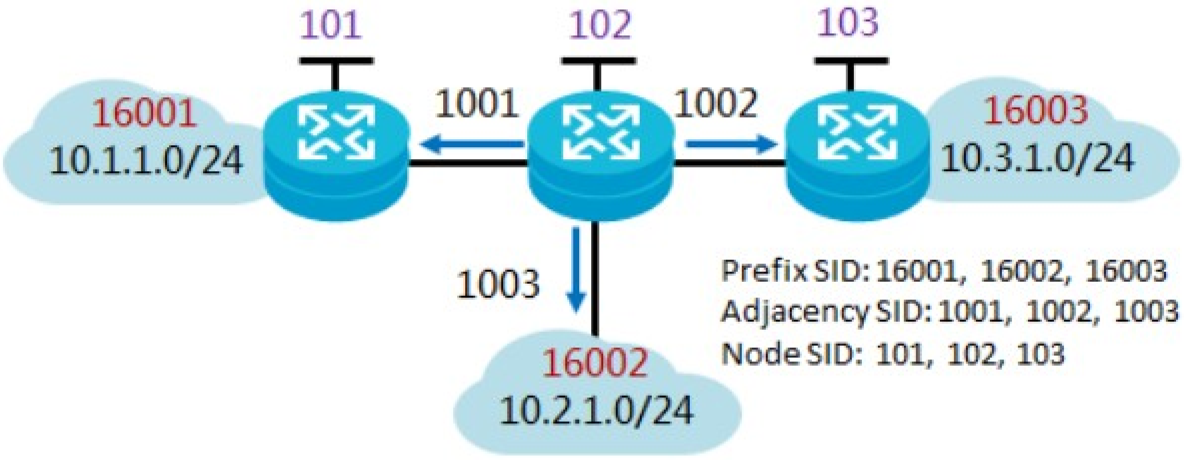
Segment Routing Fundamentals

- **Adjacency segment** (Dynamically allocated by the ingress through a protocol or manually configured) - Identifies an adjacency on a network.
- An IGP propagates the adjacency segment of an NE to the other NEs. The adjacency segment is visible globally but takes effect locally.
- Each adjacency segment is identified by an adjacency SID. The adjacency SID is a local SID that is outside of the SRGB range.

Segment Routing Fundamentals

- **Node segment** (Manually configured) - Identifies a specific node.
- Node segments are special prefix segments. When an IP address is configured as a prefix for a loopback interface of a node, the prefix SID is the node SID.

Segment Routing Fundamentals

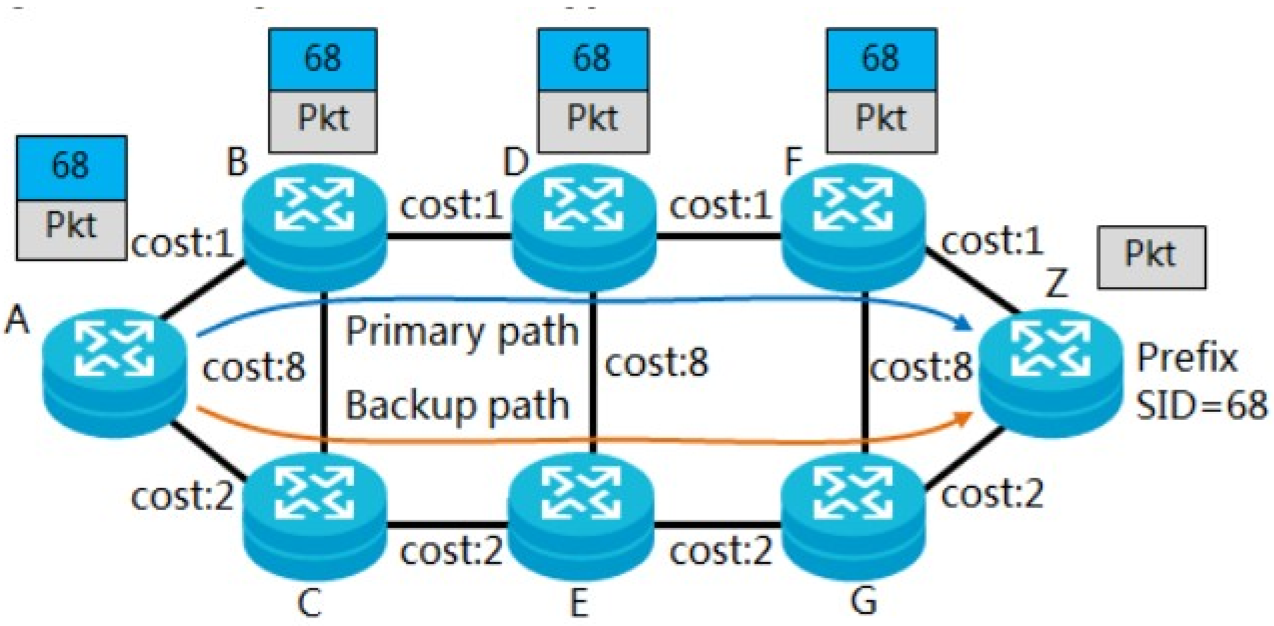


Segment Routing Fundamentals

- prefix segment indicates a destination address
- adjacency segment indicates a link for outgoing data packets
- The prefix and adjacency segments are similar to the destination IP address and outbound interface in conventional IP forwarding, respectively.

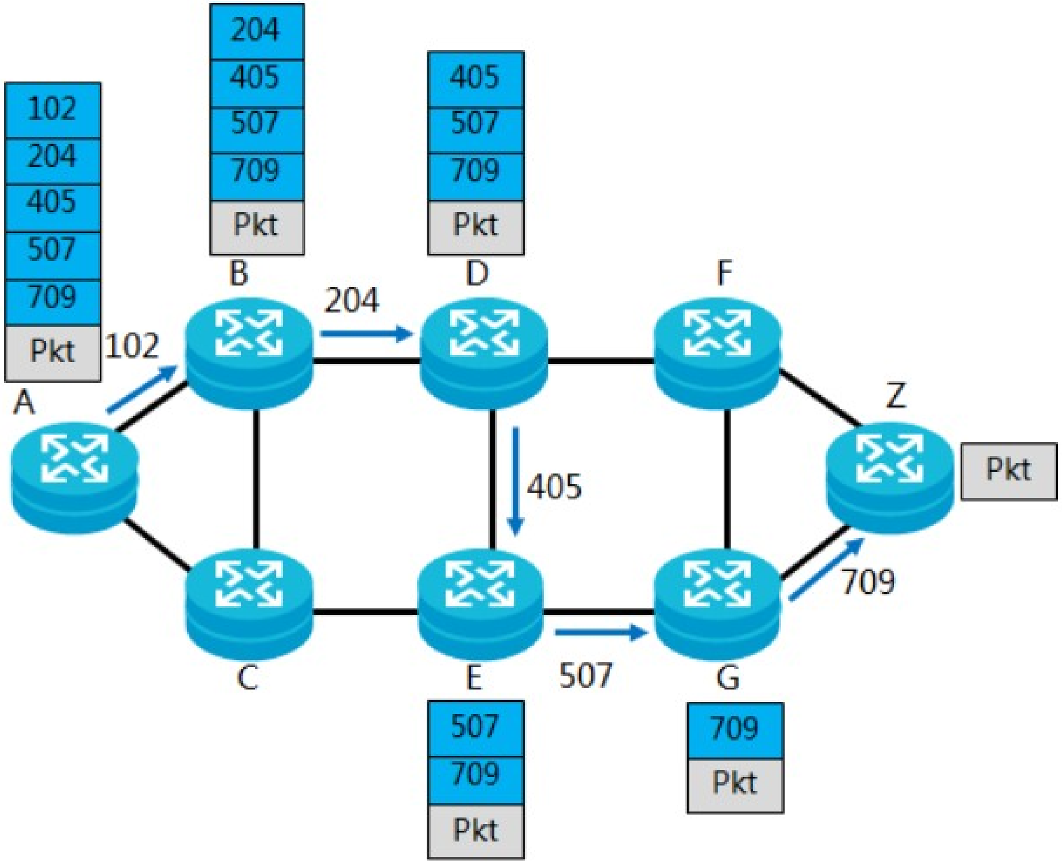
Segment Routing Fundamentals

- Prefix, adjacency, and node segments can be used independently or in combinations. They are mainly used in the following three modes:
- Prefix segment-based mode: An IGP uses the shortest path first (SPF) algorithm to compute the shortest path. This mode is also called Segment Routing-Best Effort (SR-BE).



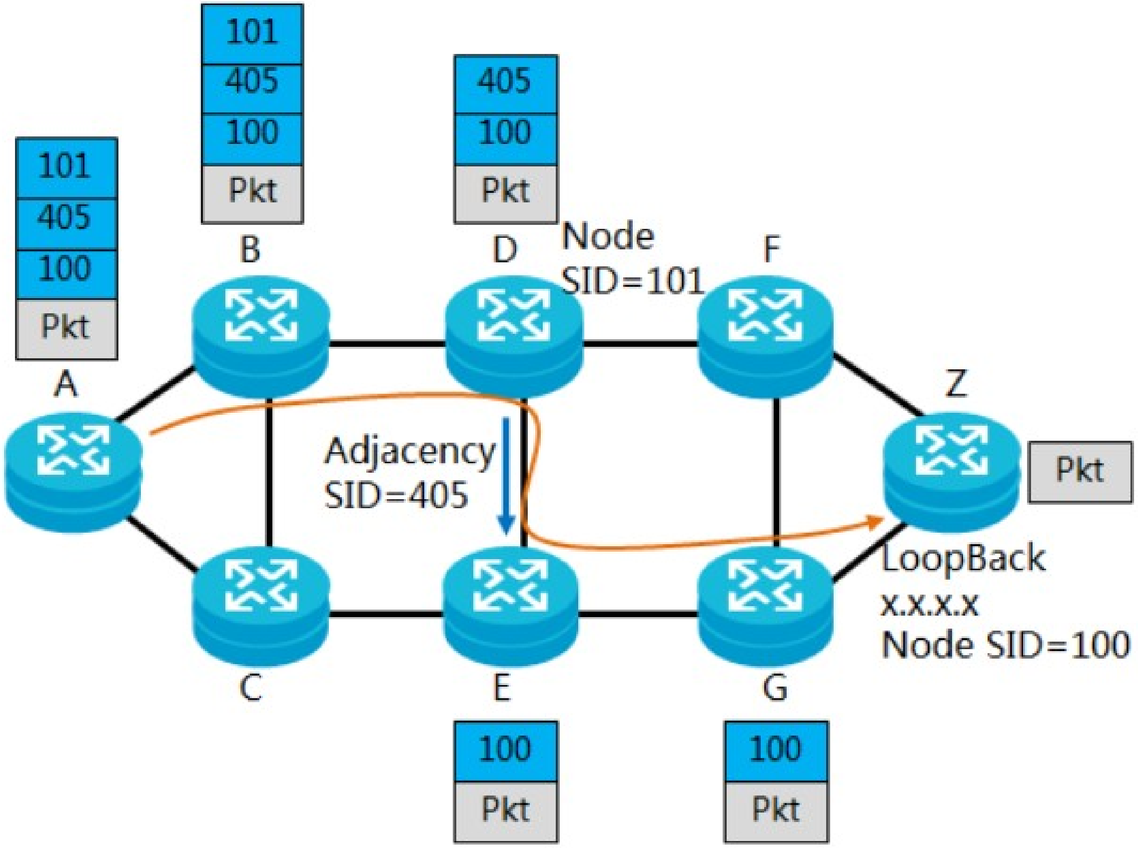
Adjacency segment-based mode

- Adjacency segment-based mode: an adjacency segment is allocated to each adjacency on the network, and a segment list with multiple adjacency segments is defined on the ingress, so that any strict explicit path can be specified. This mode is mainly used for Segment Routing-Traffic Engineering (SR-TE).



Combined mode

- Mode in which adjacency and node segments are combined: adjacency and node segments are combined, and the adjacency segment allows a path to forcibly include a specified adjacency. Nodes can run SPF to compute the shortest path based on node segments or establish multiple paths to load-balance traffic. The paths computed in this mode are not strictly fixed. Therefore, they are also called loose explicit paths. This mode is mainly used for SR-TE.



SR-BE Tunnel Establishment

- A forwarding path established using SR-BE technology is an LSP without a tunnel interface.
- This type of LSP is called SR LSP for short. The establishment and data forwarding of SR
- LSPs are similar to those of LDP LSPs.

SR-BE Tunnel Establishment

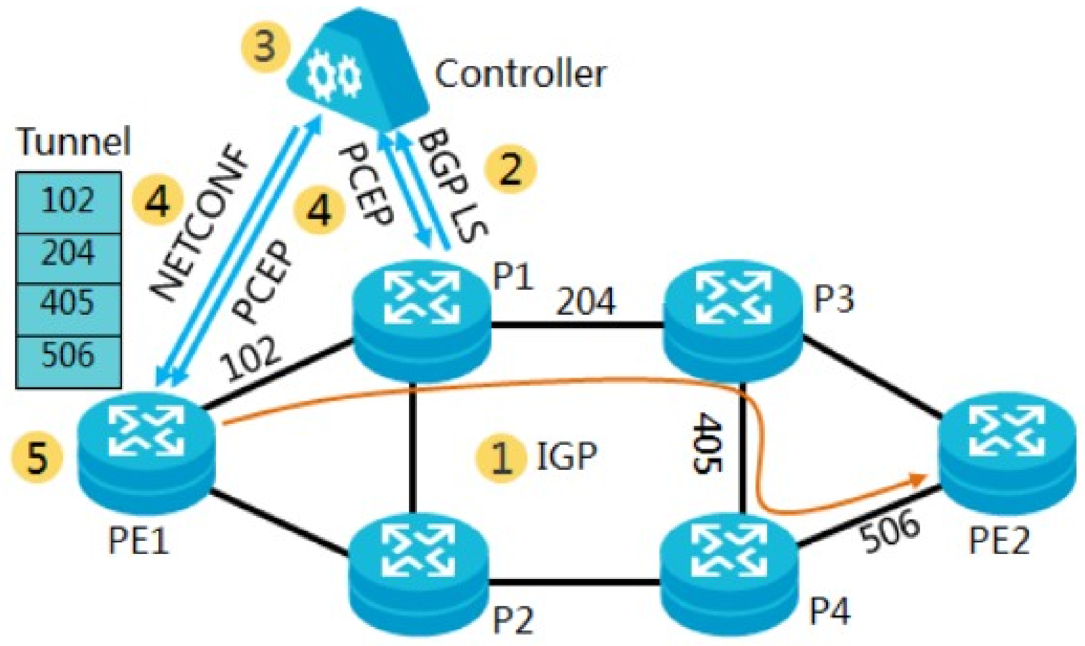
- Manual configuration: The prefix SID and SRGB are manually configured on the desired NE and then propagated through an IGP packet.
- Label distribution: Each NE parses the received IGP packet and computes a label value by summing up the start value in the local SRGB range and the prefix SID. In addition, each NE computes an outgoing label value by summing up the start value in the next-hop SRGB range and the prefix SID.
- Path computation: Based on IGP-collected topology information, the NEs use the same SPF algorithm to compute a label forwarding path and then generate a forwarding entry. Similar to traffic forwarding over MPLS LDP LSPs, traffic forwarding over SR LSPs also involves push, swap, and pop operations on label stacks and supports penultimate hop popping (PHP), MPLS QoS, and other features.

SR-TE Tunnel Establishment

- SR-TE is a new TE tunnel technology that uses Segment Routing as a control protocol. A tunnel established using SR-TE is called an SR-TE tunnel.
- SR-TE tunnels support the attributes of MPLS TE tunnels. In addition, they support
- bidirectional forwarding detection (BFD).
- SR-TE tunnels can be manually configured. Manual configuration is suitable for small-scale networks because it does not require the cooperation of a controller. However, this method does not support bandwidth reservation. In addition to manual configuration, another way to generate an SR-TE tunnel is to run the Constrained Shortest Path First (CSPF) algorithm for path computation on the ingress. Although this way supports bandwidth reservation, the computed path is only locally optimal. SR-TE tunnels can also be generated by using a controller for path computation,

SR-TE Tunnel Establishment

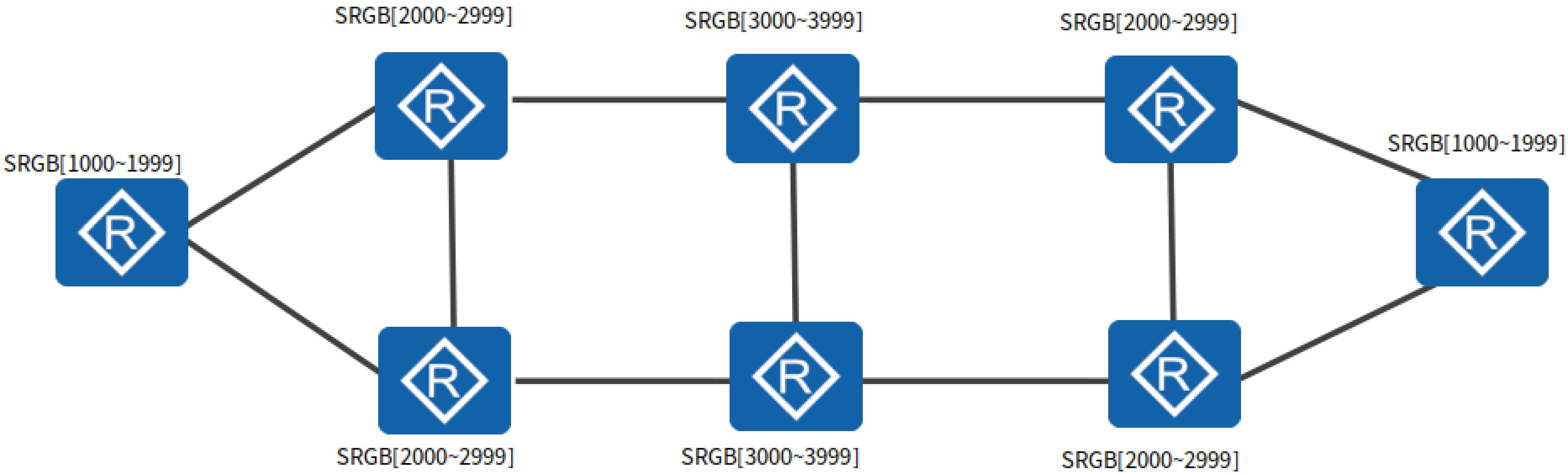
The establishment procedure is as follows:
1. Manual configuration: Configure IGP SR on forwarders to generate link topology and label information.
2. Topology and label information reporting: BGP-LS reports the information to the controller.
3. Link generation: PCEP computes a label forwarding path.
4. Information delivery: Tunnel attributes and LSP information are delivered by NETCONF and PCEP, respectively.
5. Tunnel creation: An SR-TE tunnel is automatically created between PEs based on tunnel
attributes and LSP information.



SRGB

- The SRGB is a segment range isolated from local label resources, and is specially used for Segment Routing, so that the Segment Routing global label and the traditional MPLS coexist locally.
- The node ID is the basis of the node SID. The globally unique node label must not conflict with the local label.
- The SRGB range and start value need to be configured. Based on the start value, the index value of the node SID is offset to obtain the local label.
- The scope of SRGB is not specified in the standard, and the implementation of different vendors is different.
- In the case of coexistence with traditional MPLS, the idle label space on each device cannot be the same.

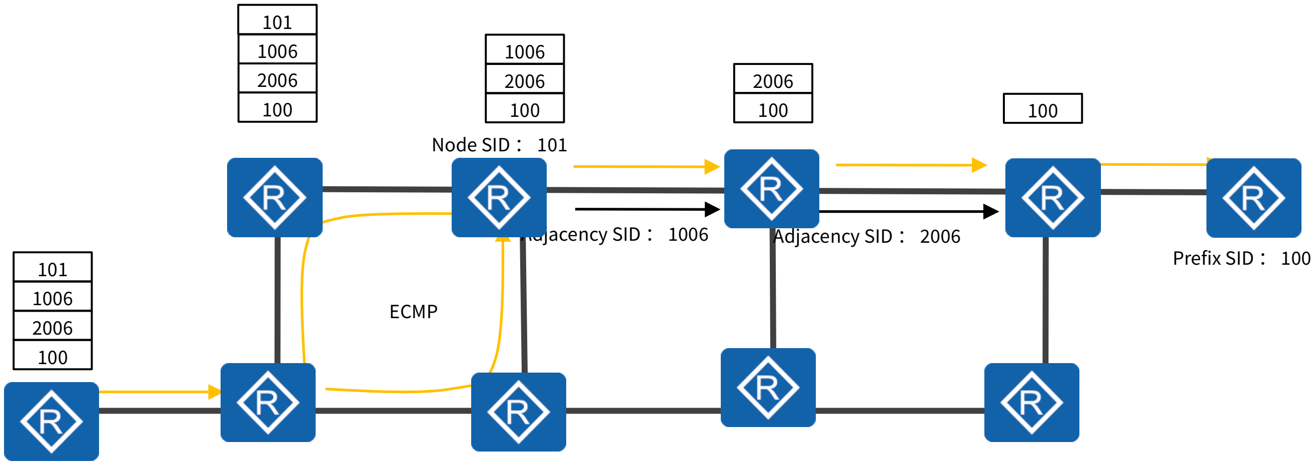
SRGB



SRGB[2000~2999]
SRGB[3000~3999]
SRGB[2000~2999]
SRGB[1000~1999]
SRGB[1000~1999]
SRGB[2000~2999]
SRGB[3000~3999]
SRGB[2000~2999]

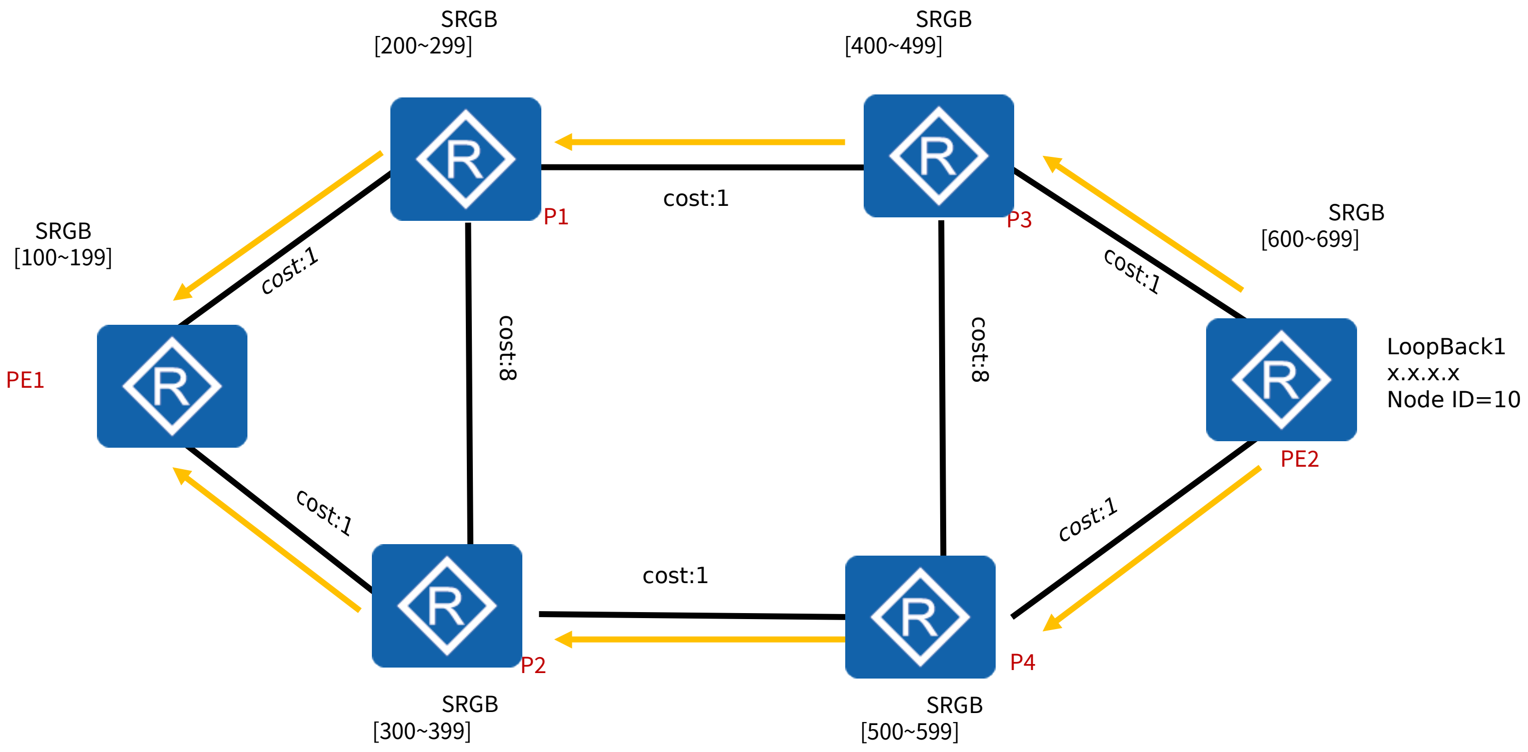
Principle of SR

SR is a protocol designed for forwarding data packets on the network based on the MPLS protocol and based on the source routing technology.



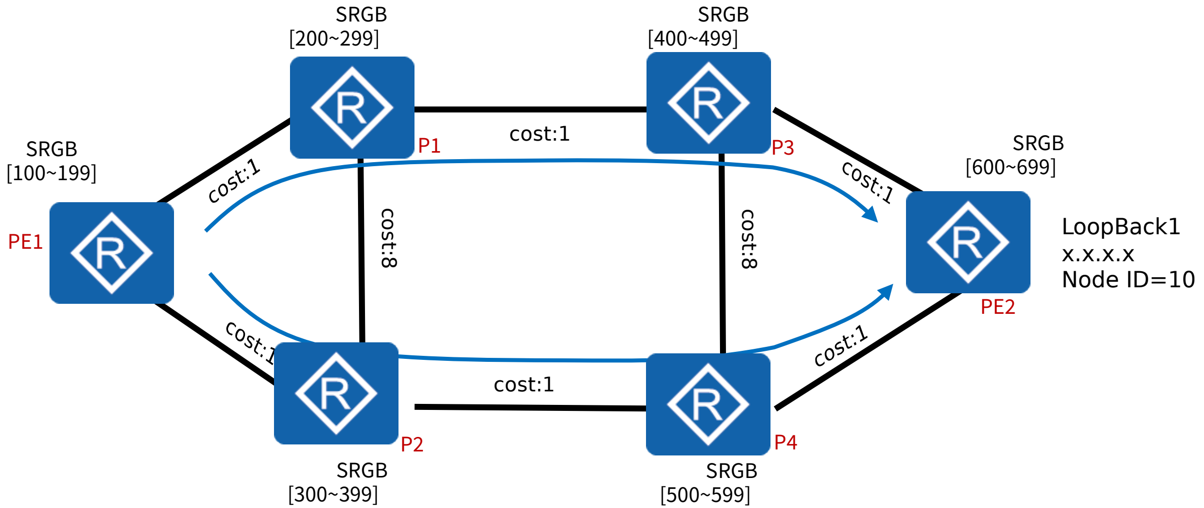
SR divides the destination address prefix / node and adjacency in the network into segments and allocates SID (Segment ID) to these segments. The Adjacency SID (adjacent segment) and Prefix/Node SID (destination address prefix / node segment) are arranged in order to obtain a forwarding path.

Working Principle of SR- (IS-IS SR-BE) (Control Plane)



Information flooding

Working Principle of SR- (IS-IS SR-BE) (Control Plane)



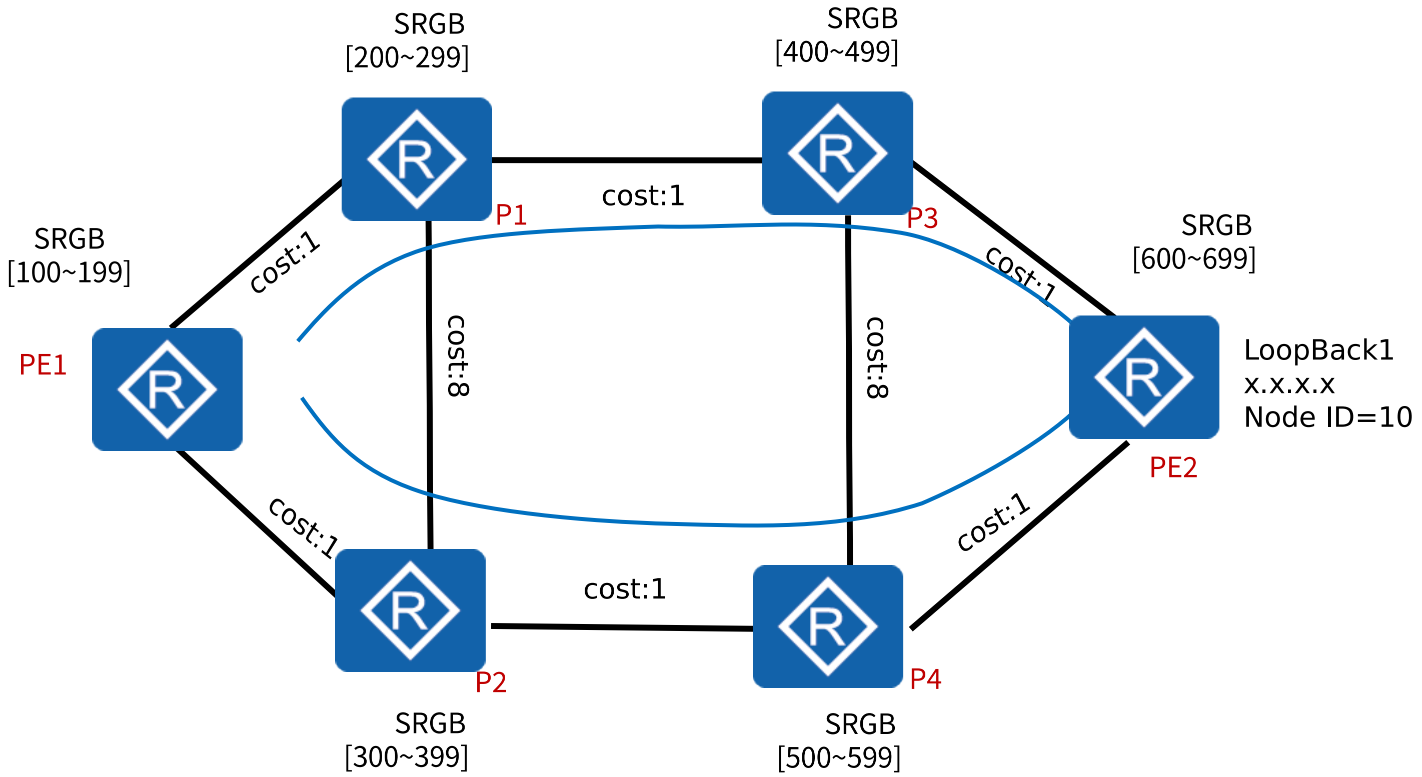
PE1→P1→P3→PE2: Label forwarding entry generated by each node for the Node SID 10

Node	InLabel	Outlabel	Interface
PE1	110	210	PE1->P1
P1	210	410	P1->P3
P3	410	610	P3->PE2
PE2	610	NA	NA

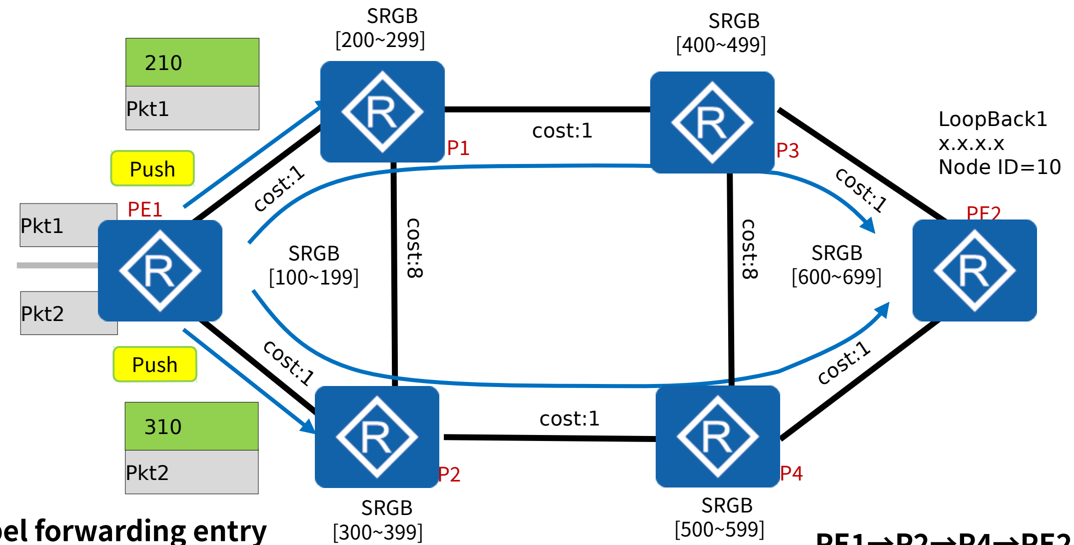
PE1→P2→P4→PE2: : Label forwarding entry generated by each node for the Node SID 10

Node	InLabel	Outlabel	Interface
PE1	110	310	PE1->P2
P2	310	510	P2->P4
P4	510	610	P4>PE2
PE2	610	NA	NA

Working Principle of SR- (IS-IS SR-BE) (Control Plane)



Working Principle of SR- (IS-IS SR-BE) (Data Plane)



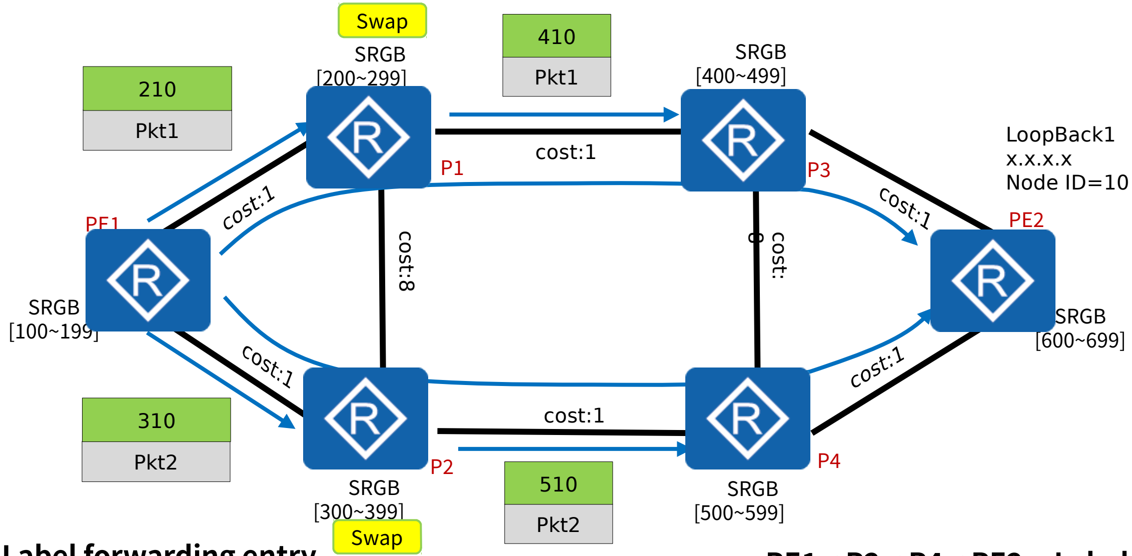
PE1→P1→P3→PE2: Label forwarding entry generated by each node for the Node SID 10

Node	InLabel	Outlabel	Interface
PE1	110	210	PE1->P1
P1	210	410	P1->P3
P3	410	610	P3->PE2
PE2	610	NA	NA

PE1→P2→P4→PE2: : Label forwarding entry generated by each node for the Node SID 10

Node	InLabel	Outlabel	Interface
PE1	110	310	PE1->P2
P2	310	510	P2->P4
P4	510	610	P4>PE2
PE2	610	NA	NA

Working Principle of SR- (IS-IS SR-BE) (Data Plane)



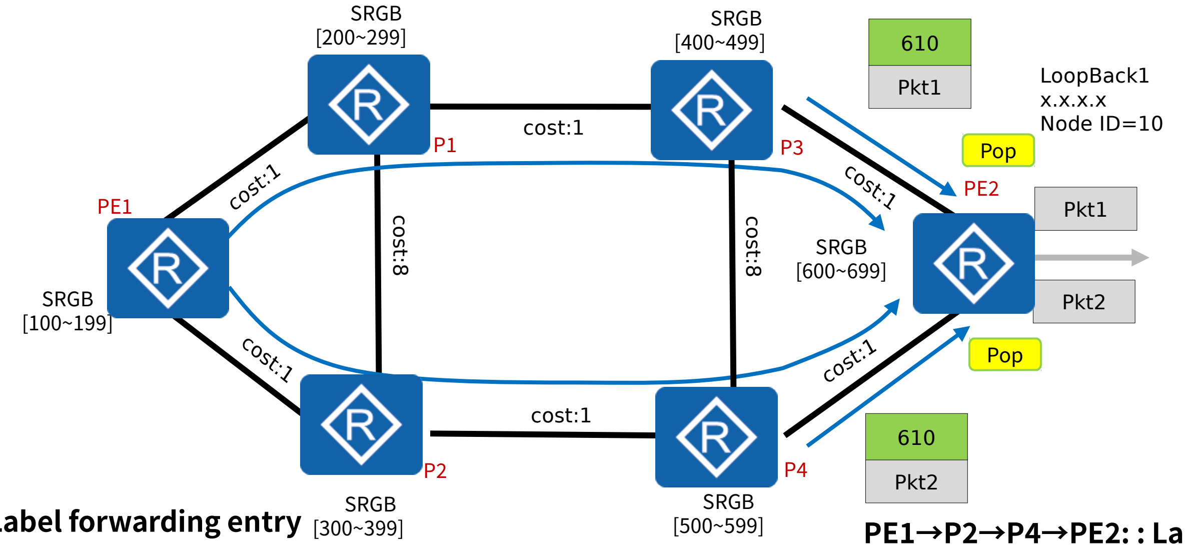
PE1→P1→P3→PE2: Label forwarding entry generated by each node for the Node SID 10

Node	InLabel	Outlabel	Interface
PE1	110	210	PE1->P1
P1	210	410	P1->P3
P3	410	610	P3->PE2
PE2	610	NA	NA

PE1→P2→P4→PE2: : Label forwarding entry generated by each node for the Node SID 10

Node	InLabel	Outlabel	Interface
PE1	110	310	PE1->P2
P2	310	510	P2->P4
P4	510	610	P4>PE2
PE2	610	NA	NA

Working Principle of SR- (IS-IS SR-BE) (Data Plane)



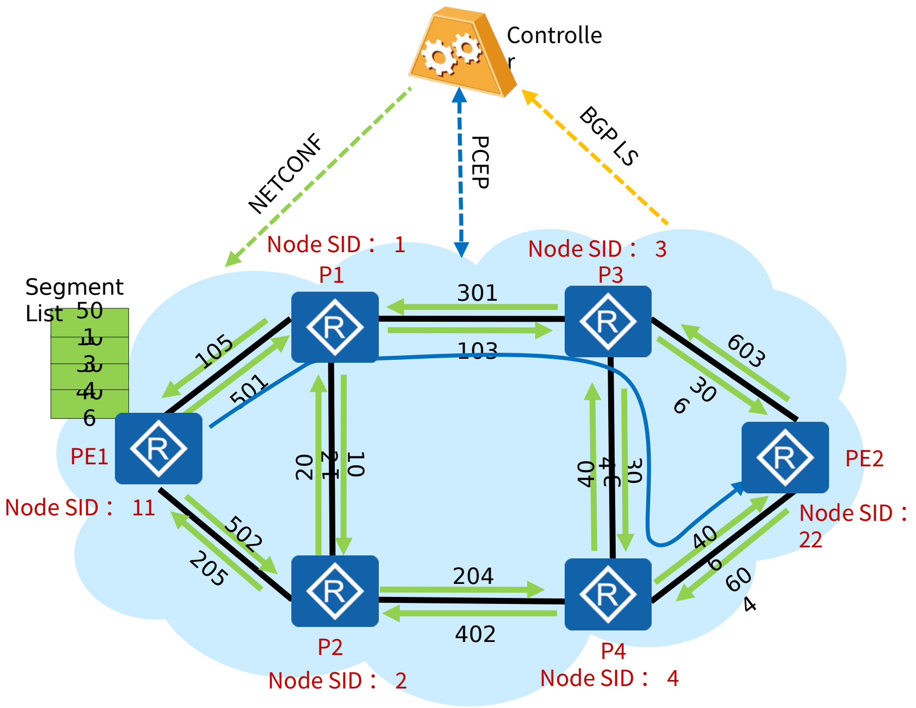
PE1→P1→P3→PE2: Label forwarding entry generated by each node for the Node SID 10

Node	InLabel	Outlabel	Interface
PE1	110	210	PE1->P1
P1	210	410	P1->P3
P3	410	610	P3->PE2
PE2	610	NA	NA

PE1→P2→P4→PE2: : Label forwarding entry generated by each node for the Node SID 10

Node	InLabel	Outlabel	Interface
PE1	110	310	PE1->P2
P2	310	510	P2->P4
P4	510	610	P4>PE2
PE2	610	NA	NA

Working Principle of SR- (IS-IS SR-TE) (Control Plane)



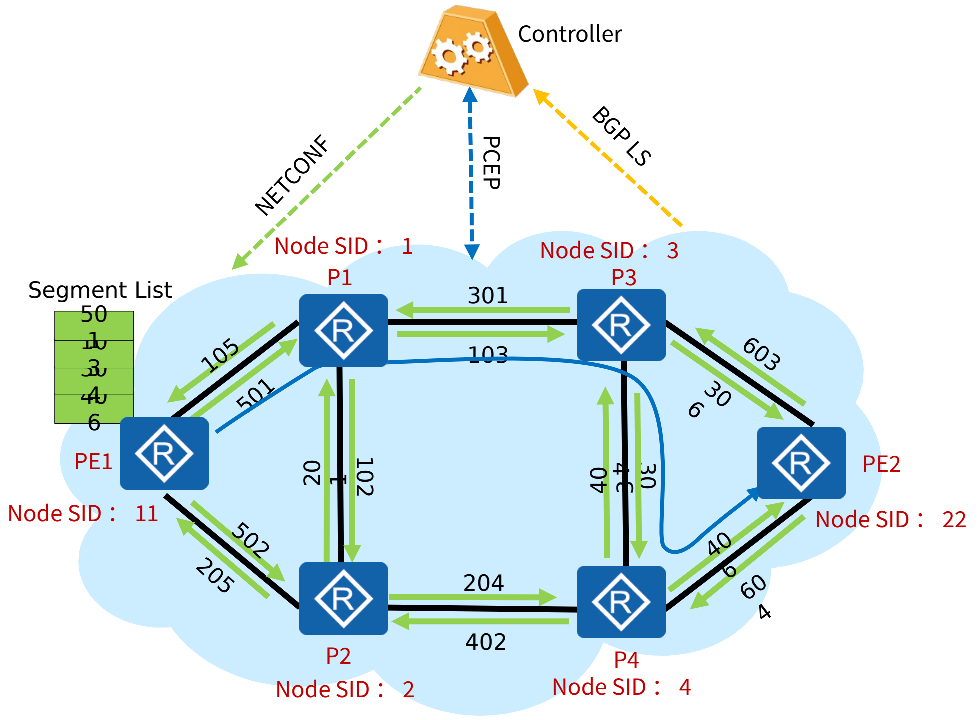
Note:

1. All nodes use the same SRGB[1000, 1999].
2. The cost values of all links are 1.

Label table on the controller

Node	Adjacency	label
PE1	NA	1011
PE1	Link PE1 to P1	501
PE1	Link PE1 to P2	502
P1	NA	1001
P1	Link P1 to PE1	105
P1	Link P1 to P2	102
P1	Link P1 to P3	103
......	......	......

Working Principle of SR- (IS-IS SR-TE) (Control Plane)



Note:

1. All nodes use the same SRGB[1000, 1999].
2. The cost values of all links are 1.

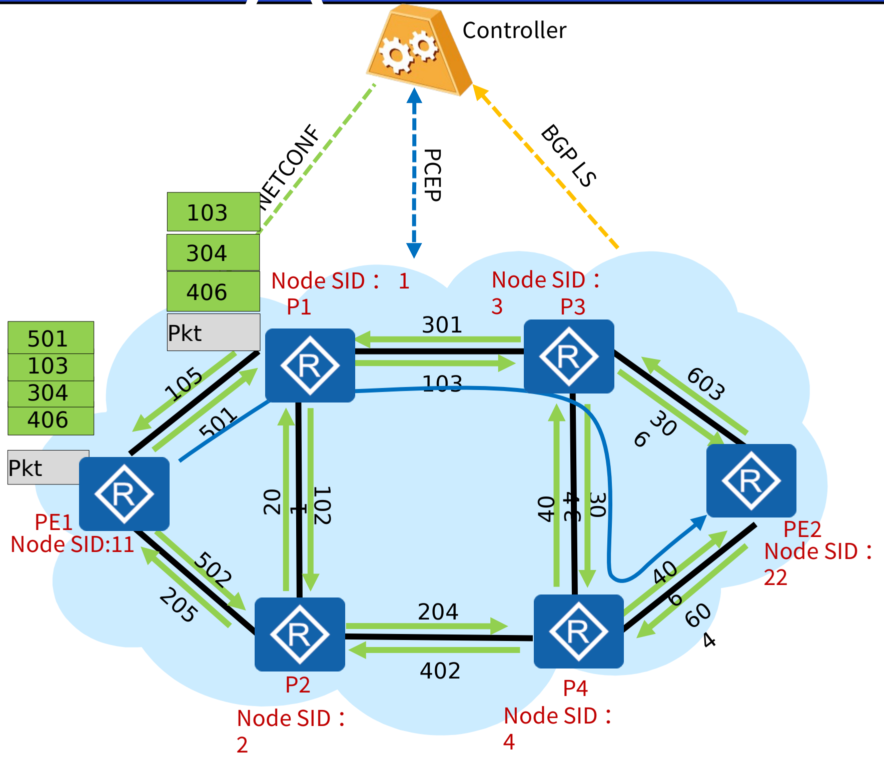
Label forwarding entries on PE1

SID/Inlabel	Outlabel	Interface
1011	NA	NA
501	NA	PE1->P1
502	NA	PE1>P2
1001	1001	PE1->P1
1002	1002	PE1->P2
1003	1003	PE1->P1
1004	1004	PE1->P2
1022	1022	PE1->P1
1022	1022	PE1->P2

SR-TE Tunnel label stack encapsulation table on PE1

SR-TE Tunnel	Segment List
Tunnel 1	501/103/304/406
......	......

Working Principle of SR- (IS-IS SR-TE) (Control Plane)



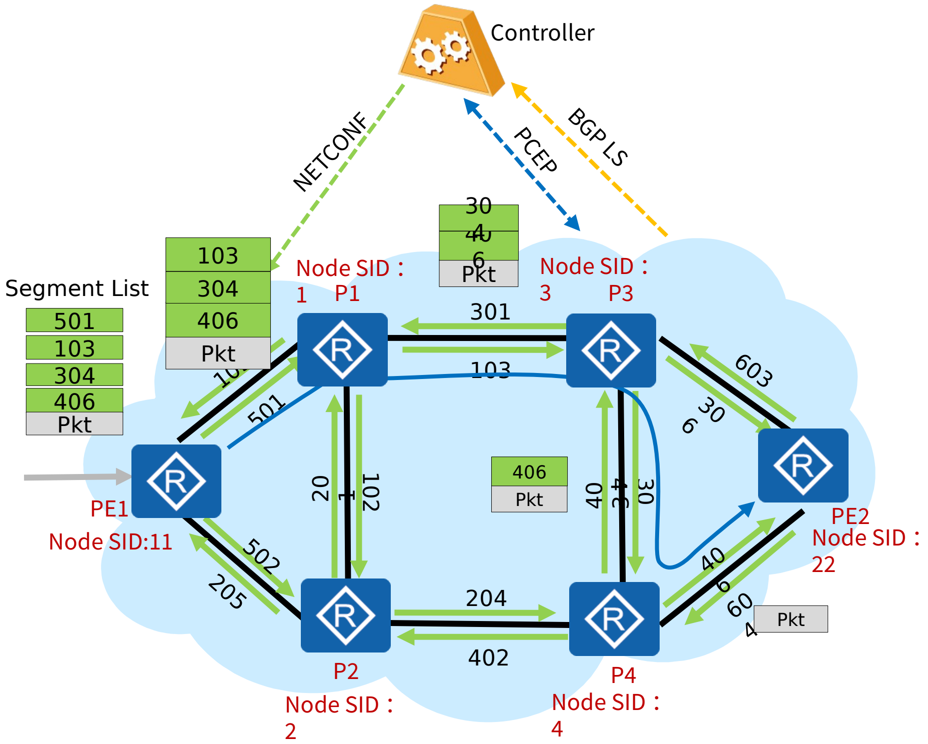
The following table shows the label forwarding entries on each node on the SR-TE Tunnel path from PE1 to PE2.

Node	SID/Inlabel	Outlabel	Interface
PE1	501	NA	PE1->P1
P1	103	NA	P1->P3
P3	304	NA	P3->P4
P4	406	NA	P4->PE2

Note:

1. All nodes use the same SRGB[1000, 1999].
2. The cost values of all links are 1.

Working Principle of SR- (SR-TE Strict Explicit Path) (Data Plane)



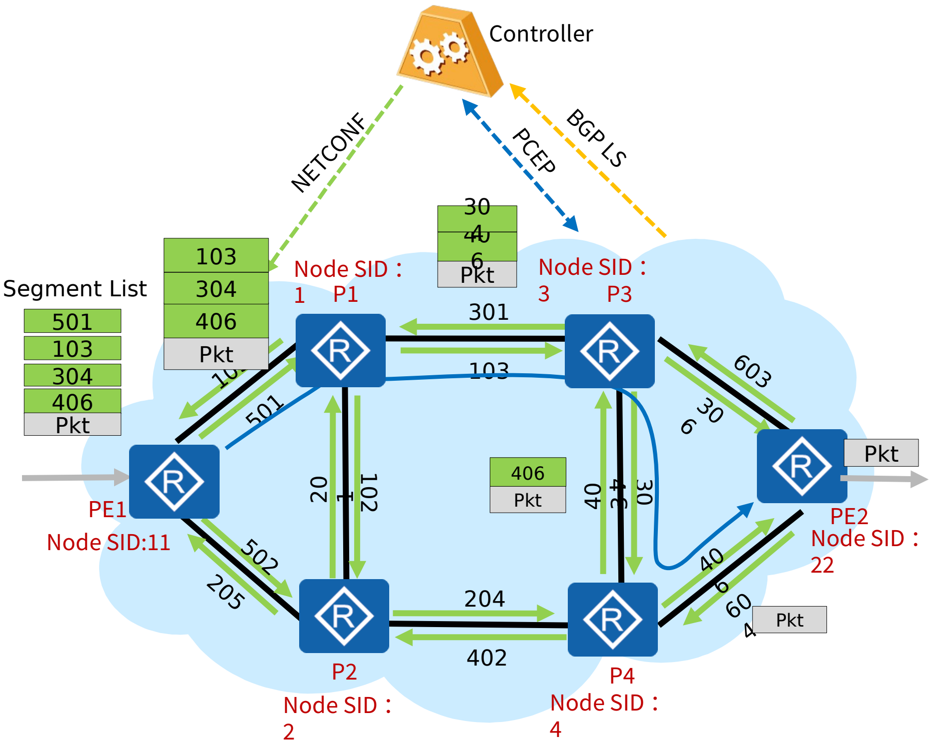
The following table shows the label forwarding entries on each node on the SR-TE Tunnel path from PE1 to PE2.

Node	SID/Inlabel	Outlabel	Interface
PE1	501	NA	PE1->P1
P1	103	NA	P1->P3
P3	304	NA	P3->P4
P4	406	NA	P4->PE2

Note:

1. All nodes use the same SRGB[1000, 1999].
2. The cost values of all links are 1.

Working Principle of SR- (SR-TE Strict Explicit Path) (Data Plane)



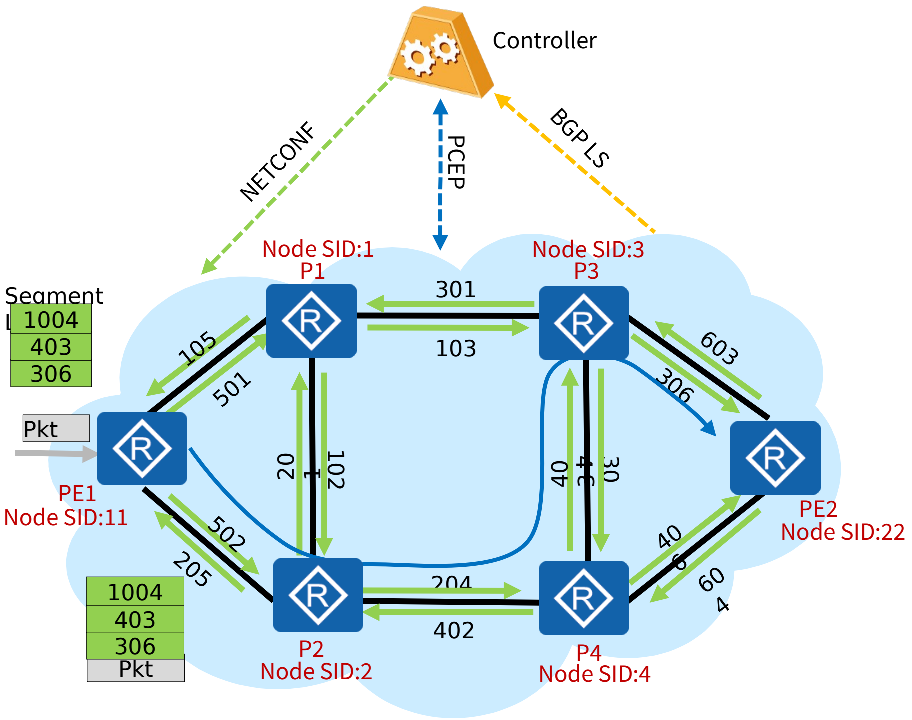
The following table shows the label forwarding entries on each node on the SR-TE Tunnel path from PE1 to PE2.

Node	SID/Inlabel	Outlabel	Interface
PE1	501	NA	PE1->P1
P1	103	NA	P1->P3
P3	304	NA	P3->P4
P4	406	NA	P4->PE2

Note:

1. All nodes use the same SRGB[1000, 1999].
2. The cost values of all links are 1.

Working Principle of SR- (SR-TE Loose Explicit Path) (Data Plane)



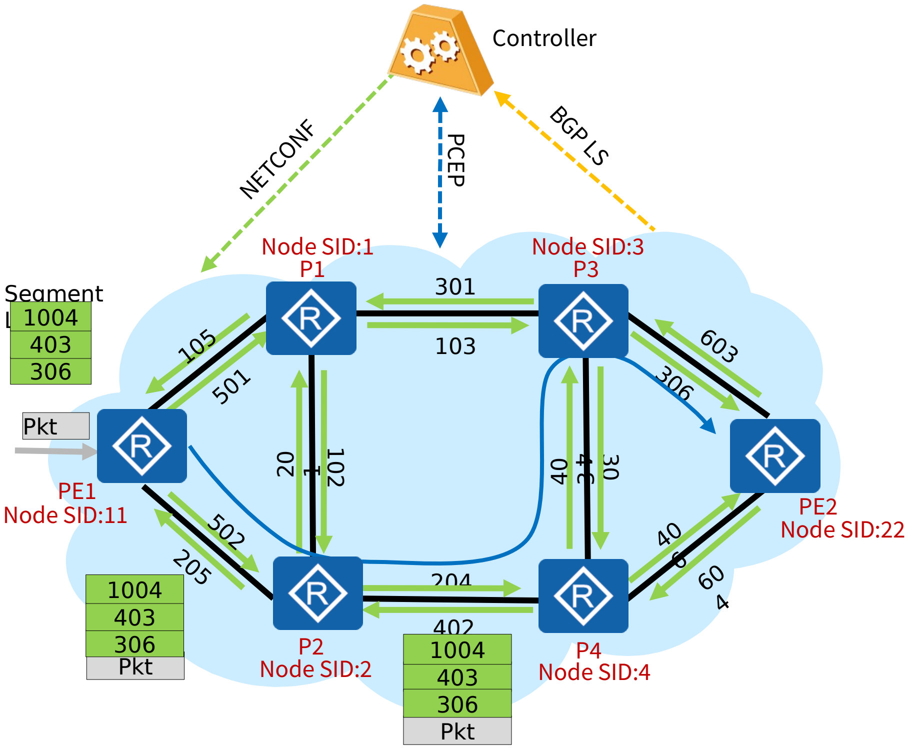
Note:

1. All nodes use the same SRGB[1000, 1999].
2. The cost values of all links are 1.

The following table shows the label forwarding entries on each node on the SR-TE Tunnel path from PE1 to PE2.

Node	SID/Inlabel	Outlabel	Interface
PE1	1004	1004	PE1->P2
P2	1004	1004	P2->P4
P4	1004	NA	NA
P4	403	NA	P4->P3
P3	306	NA	P3->PE2

Working Principle of SR- (SR-TE Loose Explicit Path) (Data Plane)



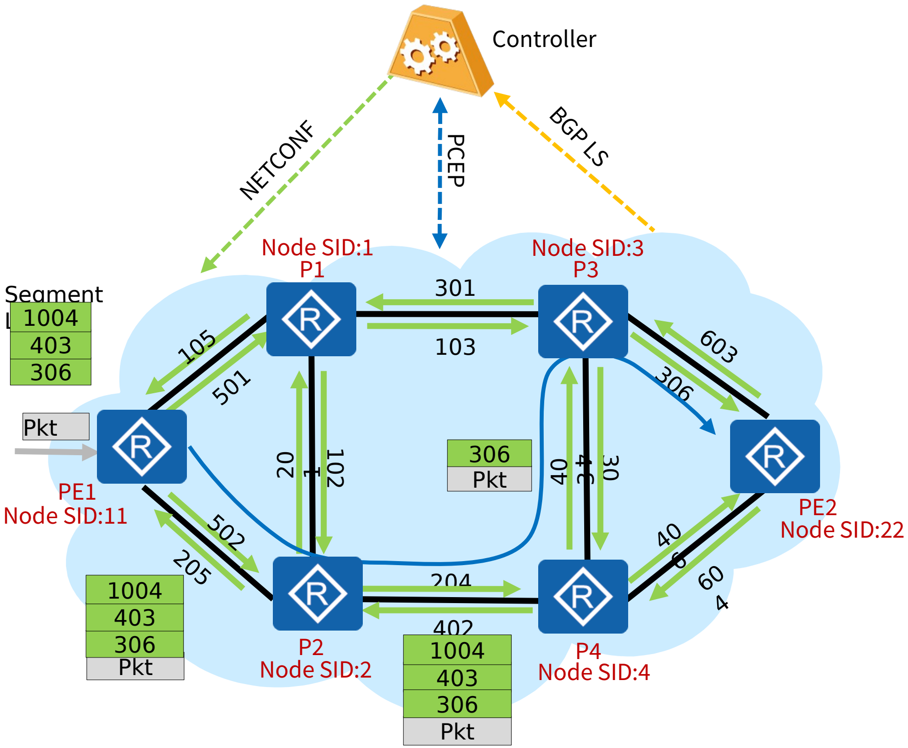
Note:

1. All nodes use the same SRGB[1000, 1999].
2. The cost values of all links are 1.

The following table shows the label forwarding entries on each node on the SR-TE Tunnel path from PE1 to PE2.

Node	SID/Inlabel	Outlabel	Interface
PE1	1004	1004	PE1->P2
P2	1004	1004	P2->P4
P4	1004	NA	NA
P4	403	NA	P4->P3
P3	306	NA	P3->PE2

Working Principle of SR- (SR-TE Loose Explicit Path) (Data Plane)



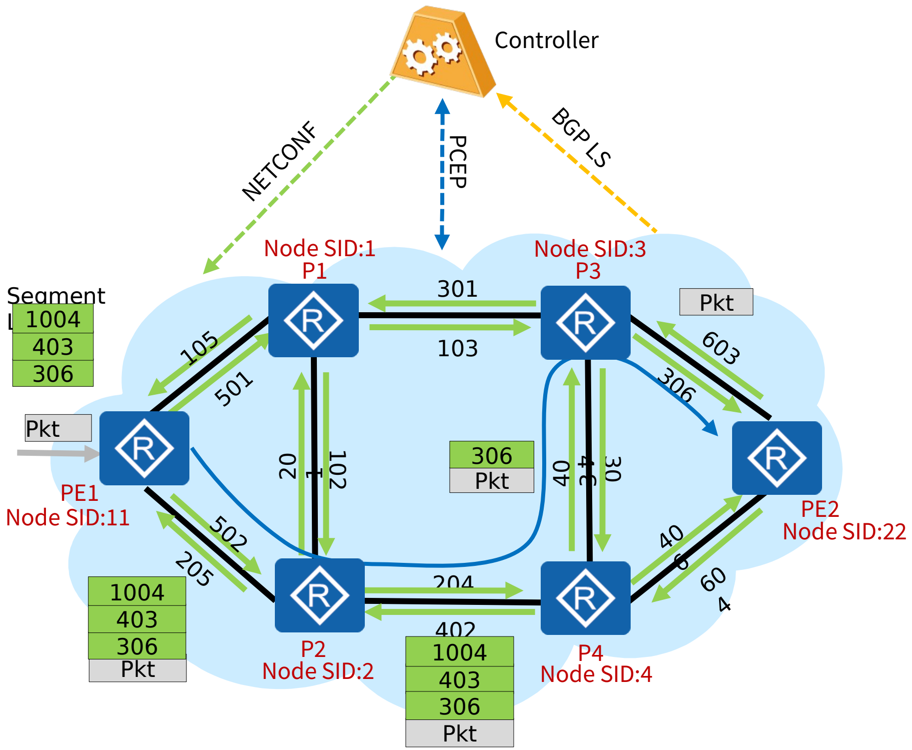
Note:

1. All nodes use the same SRGB[1000, 1999].
2. The cost values of all links are 1.

The following table shows the label forwarding entries on each node on the SR-TE Tunnel path from PE1 to PE2.

Node	SID/Inlabel	Outlabel	Interface
PE1	1004	1004	PE1->P2
P2	1004	1004	P2->P4
P4	1004	NA	NA
P4	403	NA	P4->P3
P3	306	NA	P3->PE2

Working Principle of SR- (SR-TE Loose Explicit Path) (Data Plane)



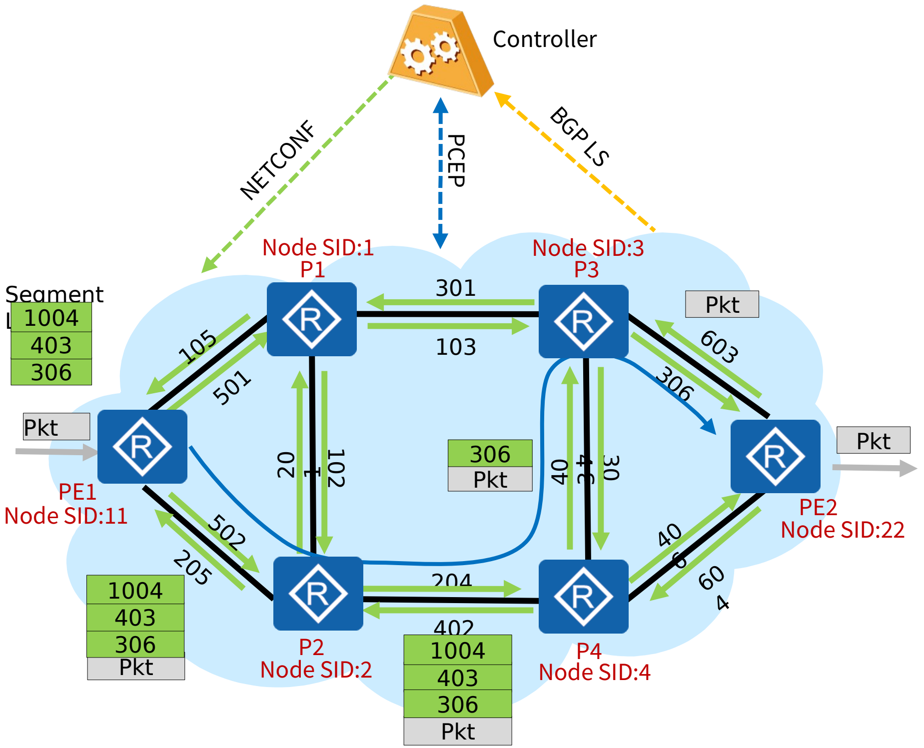
Note:

1. All nodes use the same SRGB[1000, 1999].
2. The cost values of all links are 1.

The following table shows the label forwarding entries on each node on the SR-TE Tunnel path from PE1 to PE2.

Node	SID/Inlabel	Outlabel	Interface
PE1	1004	1004	PE1->P2
P2	1004	1004	P2->P4
P4	1004	NA	NA
P4	403	NA	P4->P3
P3	306	NA	P3->PE2

Working Principle of SR- (SR-TE Loose Explicit Path) (Data Plane)



Note:

1. All nodes use the same SRGB[1000, 1999].
2. The cost values of all links are 1.

The following table shows the label forwarding entries on each node on the SR-TE Tunnel path from PE1 to PE2.

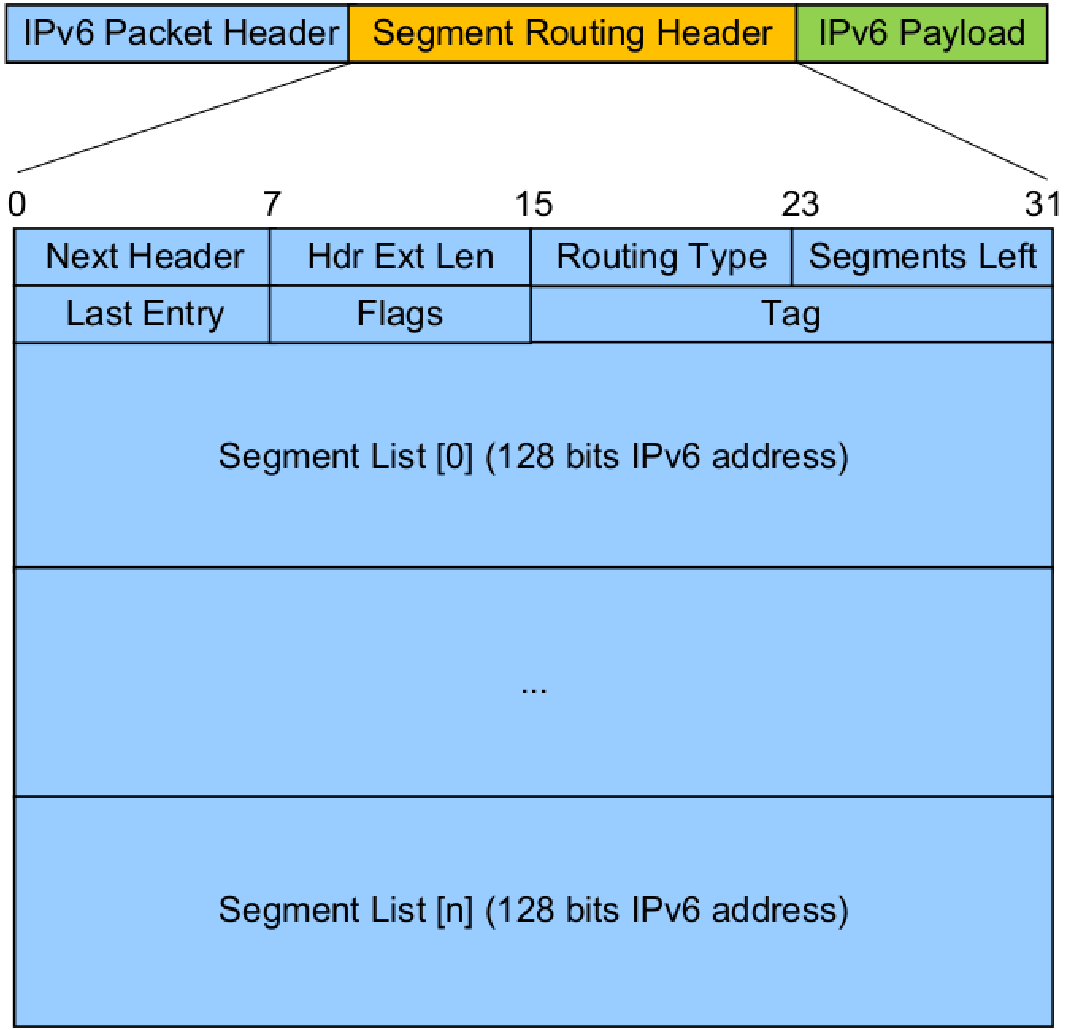
Node	SID/Inlabel	Outlabel	Interface
PE1	1004	1004	PE1->P2
P2	1004	1004	P2->P4
P4	1004	NA	NA
P4	403	NA	P4->P3
P3	306	NA	P3->PE2

Segment Routing IPv6

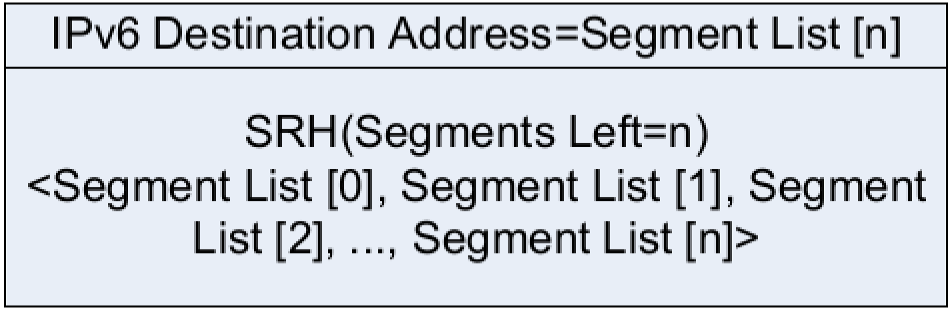
Segment Routing IPv6 (SRv6) is a protocol designed to forward IPv6 data packets on a network based on source routes. IPv6 forwarding plane-based SRv6 enables the ingress to add a segment routing header (SRH) into IPv6 packets. An explicit IPv6 address stack is pushed into the SRH. Transit nodes continue to update IPv6 destination IP addresses and offset the address stack to implement per-hop forwarding.

Segment Routing IPv6

Field Name	Length	Description
Next Header	8 bits	Header type following an SRH.
Hdr Ext Len	8 bits	SRH header length. It covers the length from Segment List [0] to Segment List [n].
Routing Type	8 bits	Route header type. Value 4 identifies an SRH type.
Segments Left	8 bits	Number of transit nodes between the existing node and the egress.
Last Entry	8 bits	Index of the last segment in a segment list.
Flags	8 bits	Identifiers in a data packet.
Tag	16 bits	Same group of packets.
Segment List[n]	128xn bits	Label segment list. A segment list is numbered from the last segment of a path. The Segment List is in the format of an IPv6 address.



Abstract SRH format



IPv6 Destination Address: IPv6 destination address in a packet. It is also called IPv6 DA. The IPv6 DA is a fixed value in an ordinary IPv6 packet. In SRv6, an IPv6 DA only identifies a next hop of an existing packet and is changeable.

<Segment List [0], Segment List [1], Segment List [2],..., Segment List [n]>: SRv6 packet segment list. Similar to an MPLS label stack in SR MPLS, it is generated on the ingress. A segment list is numbered from the last segment of a path. Segment List [0] identifies the last segment of a path; Segment List [1] identifies the penultimate segment of the path; Segment List [n] identifies the last n+1st segment, and so on. These lists are pushed into a label stack in a down-top order and numbered in an ascending order.

Segment Routing IPv6

- Each time a packet passes through an SRv6 node, the Segments Left (SL) field value decreases by 1, and the IPv6 DA changes. Both the Segments Left and Segments List fields determine IPv6 DA information.
- If the SL value is n, the IPv6 DA value is equal to the Segments List [n] value.
- ...
- If the SL value is 1, the IPv6 DA value is equal to the Segments List [1] value.
- If the SL value is 0, the IPv6 DA value is equal to the Segments List [0] value.

Segment Routing IPv6