Segment Routing Overview

Segment routing is a method of forwarding packets on the network based on the source routing paradigm. The forward path is determined before the packet is even sent. The path is encoded in the packet, at the source as a list of segments bearing forwarding instructions. At each hop, the top segment, which references the router information base (RIB), is used to identify the next hop. Segments are stacked in order, at the top of the packet header. When the top segment contains the identity of another node, the receiving node uses ECMP to move the packet to the next hop. When the identity is that of the receiving node, the node pops the top segment and performs the task required by the next segment.

Segment routing leverages other Interior Gateway Protocols such as IS-IS, OSPF, and MPLS for efficient and flexible forwarding. Segment routing is a faster and a more efficient way of forwarding traffic in the MPLS core network.

To understand the working of segment routing, let's understand how MPLS traffic engineering works.

How Does Segment Routing Work?

A router in a Segment Routing network can select either an explicit path or a default Interior Gateway Protocol (IGP) shortest path. Segments represent subpaths that a router can combine to form a complete route to a network destination. Each segment has an identifier (Segment Identifier) that is distributed throughout the network using new IGP extensions. The extensions are equally applicable to IPv4 and IPv6 control planes. Unlike the case for traditional MPLS networks, routers in a Segment Router network do not require Label Distribution Protocol (LDP) and Resource Reservation Protocol - Traffic Engineering (RSVP-TE) to allocate or signal their segment identifiers and program their forwarding information.

Each router (node) and each link (adjacency) has an associated segment identifier (SID). **Node segment identifiers** are globally unique and represent the shortest path to a router as determined by the IGP. The network administrator allocates a node ID to each router from a reserved block. On the other hand, **adjacency segment identifiers** are locally significant and represent a specific adjacency, such as egress interface, to a neighboring router. Routers automatically generate adjacency identifiers outside of the reserved block of node IDs. In an MPLS network, a segment identifier is encoded as an MPLS label stack entry. Segment IDs direct the data along a specified path. There are two kinds of segment IDS:

- Prefix SID—A segment ID that contains an IP address prefix calculated by an IGP in the service provider core network. Prefix SIDs are globally unique. A node SID is a special form of prefix SID that contains

the loop-back address of the node as the prefix. It is advertised as an index into the node-specific SR Global Block or SRGB.

- Adjacency SID—A segment ID that contains an advertising router's adjacency to a neighbor. An adjacency SID is a link between two routers. Since the adjacency SID is relative to a specific router, it is locally unique.

Examples for Segment Routing

The following figure illustrates an MPLS network with five routers using Segment Routing, IS-IS, a label range of 100 to 199 for node IDs, and 200 and higher for adjacency IDs. IS-IS would distribute IP prefix reachability alongside segment ID (the MPLS label) across the network.

*Figure 1: An MPLS Network with Five Routers Using Segment Routing*
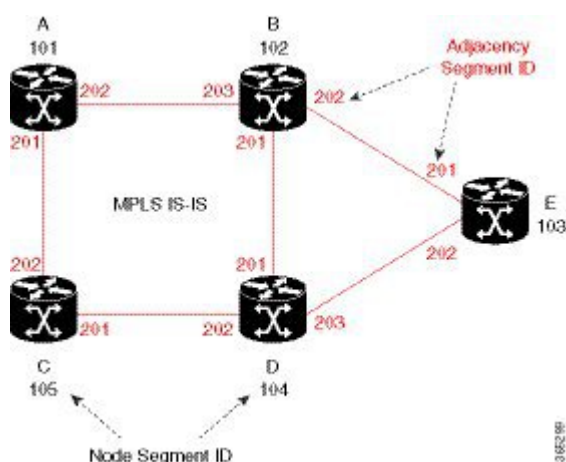
In the previous example, any router sending traffic to router E would push label 103 (router E node segment identifier) to forward traffic using the IS-IS shortest path. The MPLS label-swapping operation at each hop preserves label 103 until the packet arrives at E (Figure 2). On the other hand, adjacency segments behave differently. For example, if a packet arrives at Router D with a top-of-stack MPLS label of 203 (D-to-E adjacency segment identifier), Router D pops the label and forward the traffic to Router E.

Figure 2: MPLS Label-Swapping Operation
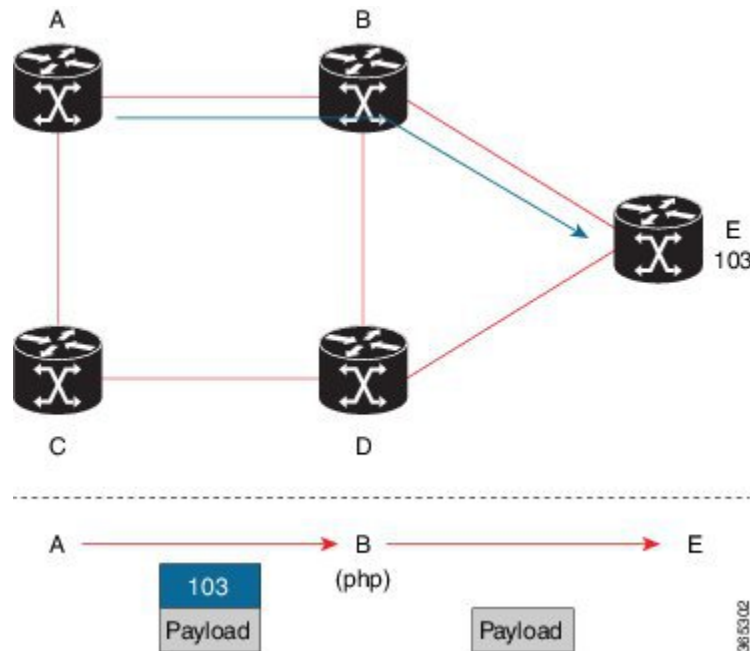
Segment identifiers can be combined as an ordered list to perform traffic engineering. A segment list can contain several adjacency segments, several node segments, or a combination of both depending on the forwarding requirements. In the previous example, Router A can alternatively push label stack (104, 203) to reach Router E using the shortest path and all applicable ECMPs to Router D, and then through an explicit interface onto the destination (Figure 3). Router A does not need to signal the new path, and the state information remains constant in the network. Router A ultimately enforces a forwarding policy that determines which flows destined to router E are switched through a particular path.

Figure 3: Destination Path of Router E
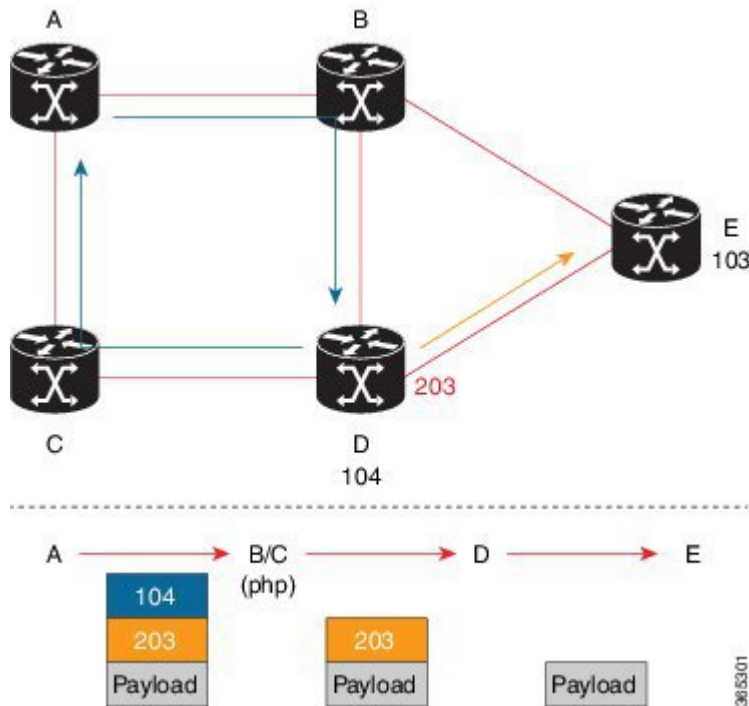
Benefits of Segment Routing

- Ready for SDN—Segment Routing is a compelling architecture conceived to embrace Software-Defined Network (SDN) and is the foundation for Application Engineered Routing (AER). It strikes a balance between network-based distributed intelligence, such as automatic link and node protection, and controller-based centralized intelligence, such as traffic optimization. It can provide strict network performance guarantees, efficient use of network resources, and high scalability for application-based transactions. The network uses minimal state information to meet these requirements. Segment routing can be easily integrated with a controller-based SDN architecture. The following figure illustrates a sample SDN scenario where the controller performs centralized optimization, including bandwidth admission control. In this scenario, the controller has a complete picture of the network topology and flows. A router can request a path to a destination with certain characteristics, for example, delay, bandwidth, diversity. The controller computes an optimal path and returns the corresponding segment list, such as an MPLS label stack, to the requesting router. At that point, the router can inject traffic with the segment list without any additional signaling in the network.

Figure 4: SDN Controller
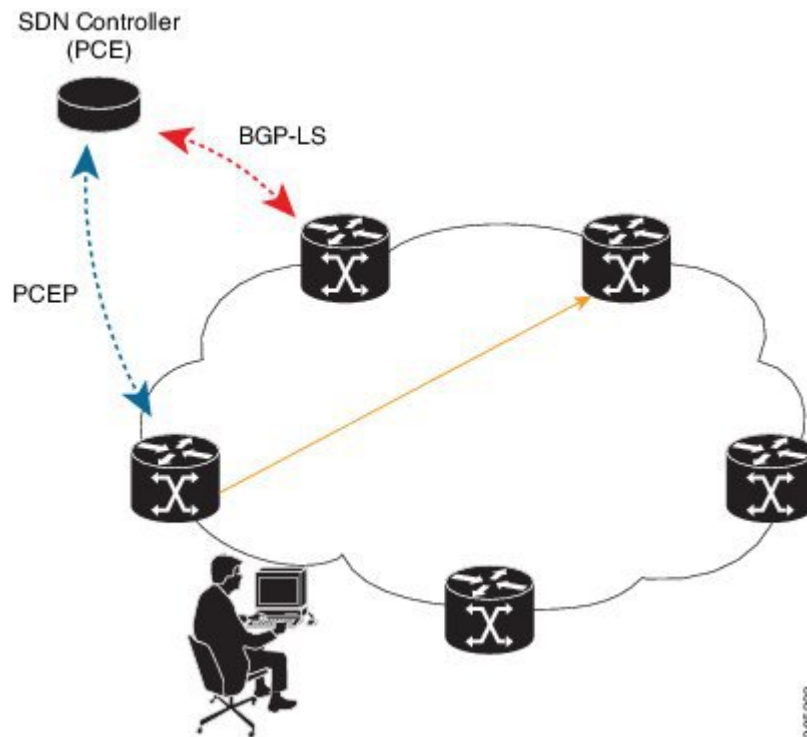
- In addition, segment lists allow complete network virtualization without adding any application state to the network. The state is encoded in the packet as a list of segments. Because the network only maintains segment state, it can support a large number - and a higher frequency - of transaction-based application requests without creating any burden on the network.

- Simplified—

    - When applied to the MPLS data plane, Segment Routing offers the ability to tunnel MPLS services (VPN, VPLS, and VPWS) from an ingress provider edge to an egress provider edge without any other protocol than an IGP (IS-IS or OSPF).

    - Simpler operation without separate protocols for label distribution (for example, no LDP or RSVP).

    - No complex LDP or IGP synchronization to troubleshoot.

    - Better utilization of installed infrastructure, for lower capital expenditures (CapEx), with ECMP-aware shortest path forwarding (using node segment IDs).

- Supports Fast Reroute (FRR)—Deliver automated FRR for any topology. If the link or node fails in a network, MPLS uses the FRR mechanism for convergence. With segment routing, the convergence time is sub-50-msec.

- Large-scale Data Center—

    - Segment Routing simplifies MPLS-enabled data center designs using Border Gateway Protocol (BGP) RFC 3107 - IPv4 labeled unicast among Top-of-the-Rack/Leaf/Spine switches.

    - BGP distributes the node segment ID, equivalent to IGP node SID.

- Any node within the topology allocates the same BGP segment for the same switch.

- The same benefits are provided as for IGP node SID: ECMP and automated FRR (BGP PIC (Prefix Independent Convergence).

- This is a building block for traffic engineering - SR TE data center fabric optimization.

- Scalable—

  - Avoid thousands of labels in LDP database.

  - Avoid thousands of MPLS Traffic Engineering LSPs in the network.

  - Avoid thousands of tunnels to configure.

- Dual-plane Networks—

  - Segment Routing provides a simple solution for disjointness enforcement within a so-called "dual-plane" network, where the route to an edge destination from a given plane stays within the plane unless the plane is partitioned.

  - An additional SID "anycast" segment ID allows the expression of macro policies such as: "Flow 1 injected in node A toward node Z must go through plane 1" and "Flow 2 injected in node A toward node Z must go through plane 2."

- Centralized Traffic Engineering—

  - Controllers and orchestration platforms can interact with Segment Routing traffic engineering for centralized optimization, such as WAN optimization.

  - Network changes such as congestion can trigger an application to optimize (recompute) the placement of segment routing traffic engineering tunnels.

  - Segment Routing tunnels are dynamically programmed onto the network from an orchestrator using southbound protocols like PCE.

  - Agile network programming is possible since Segment Routing tunnels do not require signaling and per-flow state at midpoints and tailend routers.

- Egress Peering Traffic Engineering (EPE)—

  - Segment Routing allows centralized EPE.

  - A controller instructs an ingress provider edge and content source to use a specific egress provider edge and specific external interface to reach a destination.

  - BGP "peering" segment IDs are used to express source-routed inter-domain paths.

  - Controllers learn BGP peering SIDs and the external topology of the egress border router through BGP Link Status (BGP-LS) EPE routes.

  - Controllers program ingress points with a desired path.

- Plug-and-play deployment—Segment routing tunnels are inter-operable with existing MPLS control and data planes and can be implemented in an existing deployment.

Segment Routing Limitations

Figure 5: Segment Routing State Flow
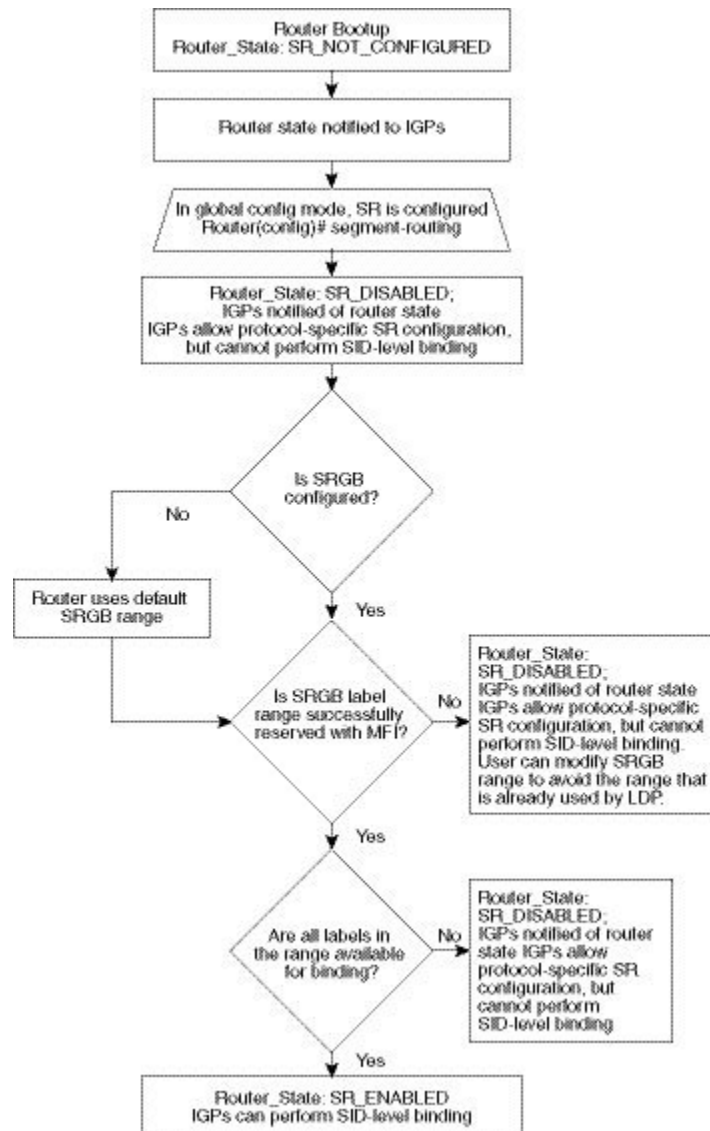
- Segment Routing must be globally enabled on the chassis *before* enabling it on the IGPs, like IS-IS or OSPF.

- Segment routing must be configured on the IS-IS instance *before* configuring a prefix SID value.

- The prefix SID value must be removed from all the interfaces under the same IS-IS instance before disabling segment routing.

General Limitations of Segment Routing

- Segment routing is supported on the Cisco ASR 900 with RSP2 and RSP3 modules.

- The Cisco ASR 900 router with RSP2 module supports five label stacks. The Cisco ASR 900 router with RSP3 module supports four label stacks. The label stacks are applicable for the default SDM template.

- The following types of services are supported:

  - VPLS

  - Layer 2 VPN

  - Layer 3 VPN

  - Global Prefixes

- To cater to scaled services, ensure that you use the following values:

  - 1500 IGP

  - 4000 L3VPN

  - 2000 L2VPN virtual circuits

- Ensure that you have the micro-loop avoidance rib-update-delay under IGP configuration as 20,000 msec. The nodes on the post-convergence path may need this time to update the forwarding plane. The number of affected prefixes may also cause a delay.

- While configuring FRR, *carrier-delay down msec 0* must be configured under interface configuration.

- Cisco ASR 900 Router with RSP3 module can push a maximum of 4 MPLS labels in the egress direction. This includes service labels (L3VPN, L2VPN, 6PE, 6VPE), RFC 3107 BGP-LU label and SR labels for FRR primary/backup paths. This restriction is not applicable while using Segment Routing Traffic Engineering. For information see Segment Routing Traffic Engineering.

Configuring Segment Routing

To configure segment routing:

1. Globally enable segment routing:

```
enable
configure terminal
```

segment-routing mpls

2. Specify the range of MPLS labels to be used to instantiate the segment routing SIDs into MPLS data plane.

```
global block
16-32767
```

Note The range of MPLS labels is 16-32767 till Cisco IOS XE Bengaluru Release 17.6.1. Starting with Cisco IOS XE Cupertino Release 17.7.1, the range of MPLS labels must be 16-40960.

3. Associate SID values with local prefix values.

```
connected-prefix-sid-map
  address-family ipv4
    92.0.0.0/24 index 51 range 1
    2.2.2.2/32 index 2 range 1
```

Configuring Segment Routing on an IGP Instance

```
enable
configure terminal
segment-routing
segment-routing [area N] {mpls | disable}
```

This command enables MPLS on all interfaces and programs the MPLS labels for forwarding.

Note If the **area** keyword is specified, segment routing is enabled only on that area.

Note The **disable** keyword can be used only if the **area** keyword is specified.

Enabling Advertisement of Mapping Server Prefix Ranges

Global segment routing configuration may contain prefix-to-SID mapping entries for prefixes that are not local to the router. Each of these entries specifies a range of prefixes. Remote mapping entries can be used to find SIDs for prefixes connected to routers that do not support SR and hence not capable of advertising SIDs themselves. This capability is part of SR-LDP inter-working functionality. OSPF learns the ranges configured in the global SR configuration and advertises them in the Extended Prefix Range TLVs.

To permit an OSPF instance to advertise mapping entries configured in the global SR mode, use the following command in router mode.

Note By default, this command is disabled. That is, no mapping ranges are advertised by OSPF even if they are configures in the global SR mode.

```
segment-routing prefix-sid-map advertise-local
```

Disabling the Mapping Server

When computing SIDs for prefixes, IGPs consider the prefix ranges received from mapping servers in the network by default. However, if this functionality needs to be disabled, use the following command in router mode. If the SR-LDP feature needs to be disabled, it is done in router-mode:

```
no segment-routing prefix-sid-map receive
```

Note   This command does not affect processing of 'native' SIDs, that is, those SIDs that are advertised in the Extended Prefix TLVs by routers to whom the prefix is locally connected.