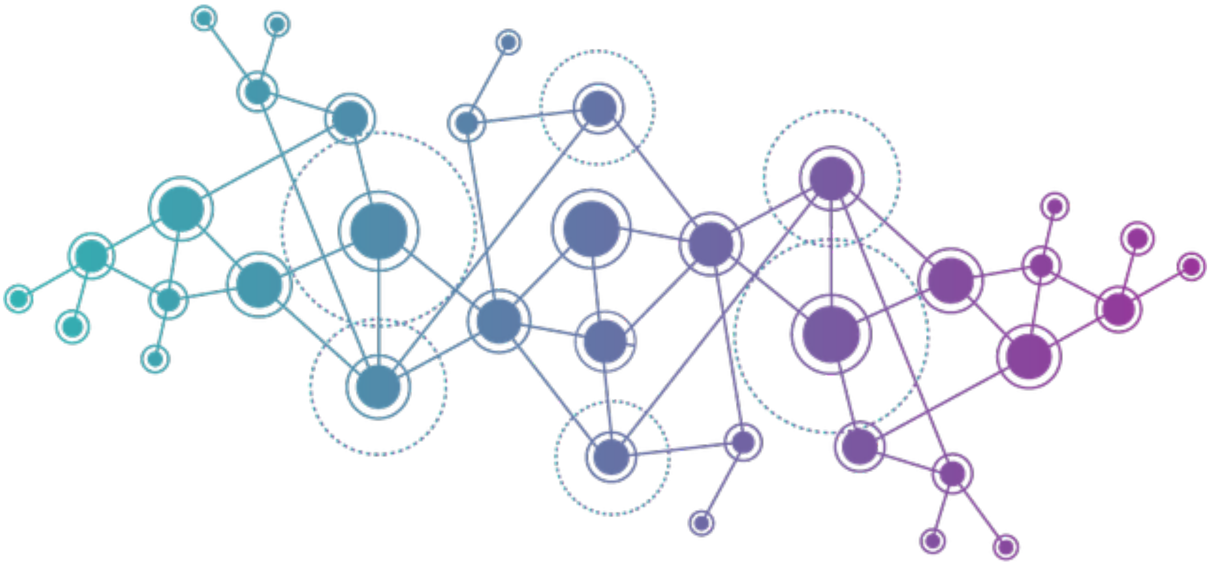


Segment Routing

Ethan Linton

Table of Contents

Purpose

The purpose of this document is to conduct an analysis surrounding segment routing. Such analysis includes why segment routing was created, how such technology operates, the technical principles involved, use cases, and the trade-offs when compared to MPLS.

The introduction of segment routing came from the many drawbacks involved when utilizing MPLS. Drawbacks within MPLS, such as the complexity presented within the control plane, the congestion due to the number of tunnels implemented, and the use of heavy-protocols such as RSVP-TE and LDP have now been fixed with the introduction of segment routing.

Design considerations that fixed issues present within previous technologies are the following - label stacks have now been replaced by segment lists, and these lists are now distributed using Interior Gateway Protocols such as IS-IS and OSPF. The need for RSVP-TE and LDP is no more, thus less complexity is present within configuration.

Use cases present within the implementation of segment routing include the likes of traffic engineering, in-which is accomplished through the use of integrating different segment identifiers (SIDs), i.e. exploitation of ADJ-SIDs to provide path avoidance. Fast Reroute can also be implemented, thus providing quick connectivity restoration after a sudden failure of network components, through the use of Topology Independent Loop-Free Alternate. Service chaining is another use case presented within this report, which covers how packets can be utilized by different services such as firewalls within the network through the use of segment identifiers.

The segment routing trade off section of this report covered a comparison between segment routing and the previous MPLS-pure networks. Such a section discussed the benefits bought upon by segment routing due to the changes introduced, such as segment distribution through IGP extensions, tunnel usage, and segment lists.

Research methods used throughout this blog included the likes of scholarly readings – as much as possible, along with readings from official network vendors such as Cisco and IETF reports.

Justification for Segment Routing

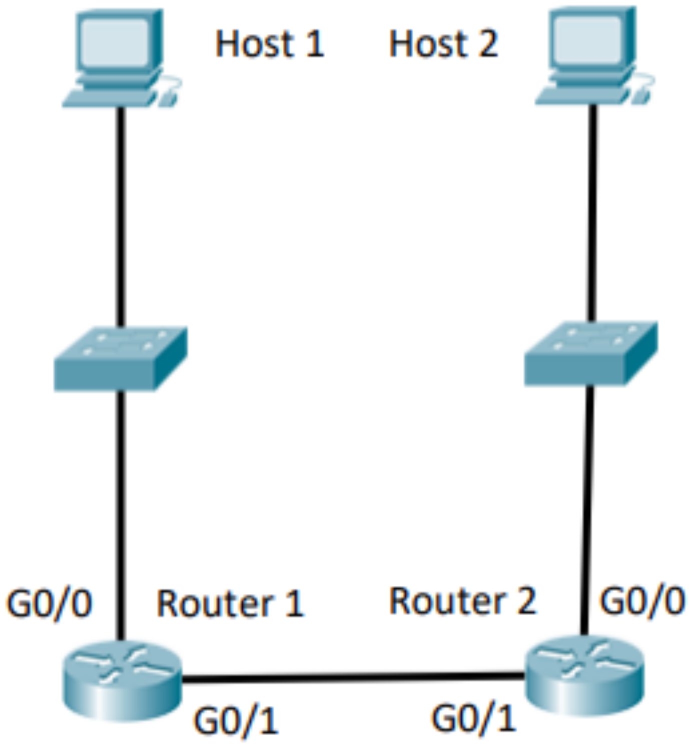
To understand the justification surrounding the implementation, and usage of Segment Routing within today's networks, the evolution from Multiprotocol Label Switching (MPLS) to segment routing must be understood, and to understand the concept of MPLS, it is imperative to understand IP routing.

IP Routing

Internet Protocol Routing, also known as IP Routing, provides a set of protocols in-which determine the paths that traversing packets follow from the packet's source to its destination through a succession of routers. In the case of IP Routing, these protocols include the likes of:

- Open-Shortest-Path-First
- Routing Information Protocol
- Border Gateway Protocol
- Intermediate System to Intermediate System [1]

The following is an example basic network topology, in-which explains how a packet is forwarded through a network using OSPF supplied by IP Routing:



Host 1	10.0.0.0
Host 2	10.1.1.0
Router 1	G0/0 - 10.0.0.0 G0/1 - 10.2.2.0
Router 2	G0/0 - 10.1.1.0 G0/1 - 10.2.2.0

1. Appropriate IP Addresses are assigned to devices.

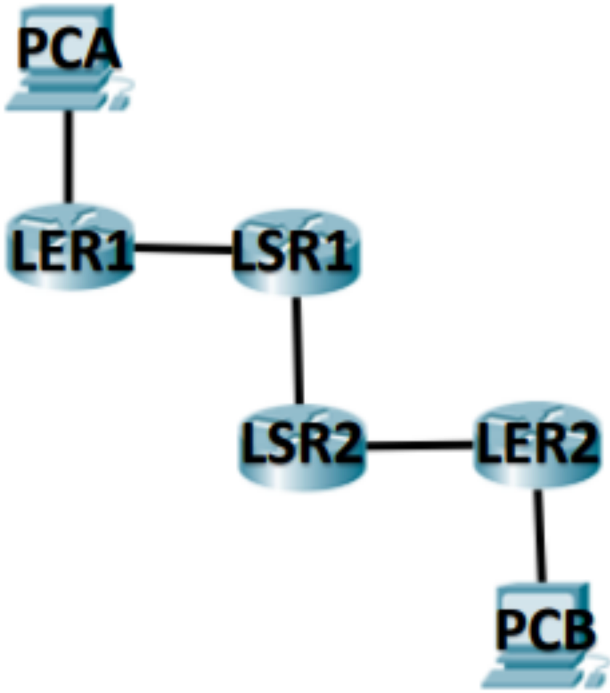
2. Each router will populate their routing table with the assigned IP addresses.
3. For example, Router 1 is not aware that 10.1.1.0 exists behind Router 2.
4. OSPF is used to advertise the networks that both routers are connected to, they fetch this information from their respective routing tables. Both Routers advertise it's interfaces and populates their routing tables with the new routing information.
5. Both Routers are now aware of the complete topology. [1]

Once OSPF has established a full network topology, packets are then forwarded through the network via identifying the correct ingress/egress interfaces that the packet shall pass using IP Lookup.

IP Routing is known to have poor flexibility, experience performance issues, and lack path control. Thus, some of these previously mentioned issues with IP routing were fixed with the creation and implementation of Multiprotocol Label Switching (MPLS). [1]

Multiprotocol Label Switching

Multiprotocol Label Switching, also known as MPLS, is a forwarding mechanism in-which utilizes the use of labels. Defined for MPLS, a label is a simple binary integer that is used to determine the outgoing interface for the next hop by the receiving device forwarding table. The process behind the usage of labels within MPLS is as follows – the ingress label is swapped with the outgoing label, then the packet is forwarded to the next device. Such process is then potentially repeated multiple times until the packet reaches its desired destination. The routing devices allocated for the use of MPLS are as follows – Label edge router and the label switch router. [1]



The label edge routers are depicted by 'LER', and the label switch routers are depicted by 'LSR'.

At both edge routers, a label forwarding table resides. This table binds labels to a destination IP address, which in this case is '10.0.0.1' which is PCB. The label assigned to this IP address is 200. In this case, the label 200 needs to be pushed onto the packet being sent from PCA, to PCB. The packet with this label is forwarded to the next hop being LSR1.

LSR1 has a label information base (LIB). This table holds information surrounding the incoming label, outgoing label and outgoing interface. Label 200 is swapped with 300, then forwarded to LSR2.

This exact process is done on LSR2, and the label is swapped. When the packet reaches LER2, the label is popped from the packet, and forwarded to its destination. [1]

In comparison to the previously mentioned IP Routing, MPLS provides more flexibility, paths are predefined with assigned labels allocated to each packet for forwarding, and traffic engineering is simple to implement. Although there is a downfall when implementing such technology, when attempting to maintain a more complex control plane when utilizing MPLS, the cost surrounding maintenance is high. The solution for addressing such noted issues is to connect all network nodes within the network to one main controller which makes the infrastructure more scalable and programmable and configure such nodes to communicate using a common protocol. Such concept leads to the creation of Segment Routing that provides more scalability, flexibility, and is simpler to operate compared to the previously noted IP routing and MPLS technologies. [1]

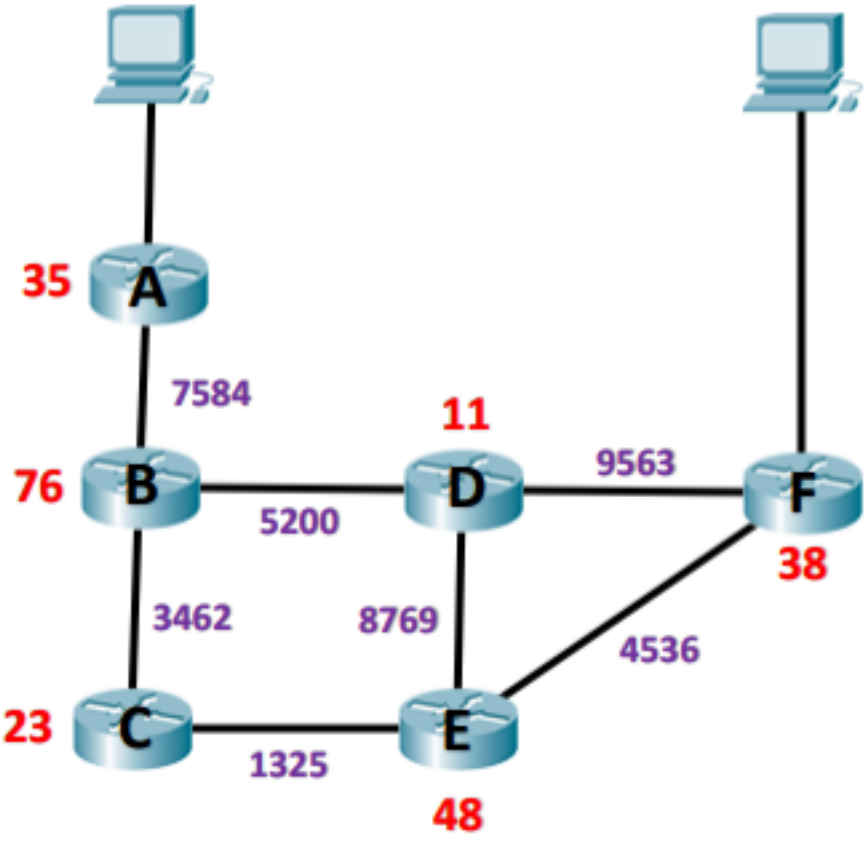
Segment Routing

The creation surrounding the segment routing architecture was based on the real world, meaning that in result of this, the use of minimal shortest path hops through intermediate devices within the network is needed. The scalability and overall functionality of the network are increased with the use of segment routing due to labels being distributed through Interior gateway protocols, and explicit paths being built using equal-cost multipath. In this particular design, the implementation surrounding the MPLS control plane will see no need for the use of RSVP-TE or LDP, thus resulting in less complexity. Such technological elements surrounding segment routing is explained in the next section. [2]

Technology Principles – Segment Routing

To understand the technological principles presented within the segment routing architecture, we will analyze the overall process in-which how segment routing is used to determine paths of a packet flow.

The following is a simple topology, with segment routing enabled.



Topology Example

- Six routers are shown within this topology as A, B, C, D, E, and F.
- A red nodal ID is allocated to each router.
- Segment IDs are established between each router within the topology, shown in purple. [1]

Once the segments IDs are established between each router within the network, they are then distributed to every router within the network through the use of an Interior Gateway Protocol. For segment routing, the use of Open-Shortest-Path-First (OSPF) or Intermediate System – Intermediate System (IS-IS) may be used as the chosen protocol for distributing segment IDs. This notes one of the main differences between segment routing and MPLS. When utilizing MPLS, labels are distributed by RSVP-TE or LDP. But, within segment routing, segments are distributed through the use of the chosen IGP configured within the routers. As noted previously, such difference grants segment routing less configuration complexity due to the decrease of additional configuration on devices. [3]

Intermediate System – Intermediate System, also known as IS-IS, is a link-state routing protocol in which may be utilized if chosen by segment routing. The similarities between the link-state routing protocols available for use within segment routing are the following –

- Both, IS-IS, and OSPF utilize the Dijkstra SPF algorithm
- Both are IGP, meaning they distribute routing information between routers that belong to the same Autonomous System (AS)
- Both have support for multi-path, authentication, variable subnet length masking, and classless inter-domain routing. [4]

When considering what link-state routing technology to implement when configuring segment routing, the overall design, and infrastructure of the network must be considered. As such, for example, the implementation of OSPF is more suited for networks that have an area 0 core, with sub-areas distributed around, also known as a rigid area design. The implementation of IS-IS is more suited for the following - when layer 2 routers are linked through the backbone,network has a diverse infrastructure, and when flexibility is crucial. In this example, OSPF will be used. [4]

Open-Shortest-Path-First is utilized within segment routing by sending hello packets to every router in the network, along with the segment IDs to the forwarding table. The result of sending and receiving the hello packets is that now each router is aware of each other. OSPF continues to flood the network surrounding information about each router, this means that OSPF maintains an entire logical picture of the topology, as such if a router were to shutdown expectedly or unexpectedly, OSPF will know about it and alert the entire network. If such situation occurs, OSPF will help the source router establish the next best shortest path. In the case of segment routing, the term ‘shortest path’ does not only mean the number of hops, but there are also additional metrics put to use, which is load balancing, exception handling, and bandwidth. We can now see that the use of an Interior Gateway Protocol in the case of segment routing is performing critical operations, such as, distributing segments IDs throughout the network, advertises the routers, calculates the shortest path, and passes the information surrounding the calculated shortest path to the forwarding routing table. [1]

The process presented within the implementation of segment routing is the following –

1. The packet is received by Router A, from the host, destined for the other host.
2. Router A, known as the ingress node, now attaches a segment routing header to the packet, which includes a segment list. The segment list includes information on how the received packet should be forwarded

to its desired location. In this example, the information included in the segment list consists of the following – 23 to 48 to 38. These numbers are the nodal IDs in-which the packet must pass through to reach its destination.

3. Router A, being the source router, examines its routing table for the next hop to reach the packet's destination, which in this case is 23, and a destination ID, being the IP address of the end-destination.
4. When the packet is forwarded to Router B, the only task in which is performed at Router B is the examination of the top entry within the included segment list. The top entry within the segment list is '23', as mentioned previously, so the packet is instantly forwarded to Router C which has a nodal ID of 23.
5. Routers C, E, and F, which have their assigned nodal IDs within the segment list use their label information base table (LIB), which contains information such as the outgoing interface for the packet, and incoming/outgoing segment ID. This presents the main benefit of such routing, routers within the network do not have to maintain information in the routing table because the forwarding instructions are within the received segment. Once the packet arrives at Router C, the label information base tables contain 23 as the incoming segment ID, 48 as outgoing segment ID, and the outgoing interface through which the label needs to be forwarded. So, the packet is forwarded to segment 48, which is Router E.
6. Once Router E receives the packet, the only segment IDs contained in the segment list are 48 and 38. As represented within Router E's LIB table, the incoming segment ID is 48, and the outgoing segment ID is 38. The segment ID 48 in which the packet was received on is popped, and the packet is only forwarded with the segment ID of 38, which is the nodal ID for Router F.
7. Once Router F receives the packet, the packet is then forwarded to the desired destination being the other host device. [1]

Segment Routing Technology Principle Explanation

Segment

As defined by RFC 8402, surrounding Segment Routing Architecture, the term 'Segment' is defined as an ordered list of instructions in-which a node transverses a packet through. [5] Such a list of instructions may include the likes of forwarding the received packet to a specific node/interface or delivering the packet to a specific service or application. A segment identifier, also known as 'SID' is used to uniquely identify each segment. [6]

Local and Group Segments

Within a segment routing domain, a segment, as explained previously, can either be of type local or global. Such local segments are applied to segment routing nodes, and global segments are applied to the segment routing domain.

Local segments are utilized within segment routing networks to only have local significance. What this means is that a router is not aware of other local segments within other routers of the same domain, it is only related to that specific local router forwarding information base (FIB). Local segments also take a value outside of the SR Global Block (SRGB) range. The segment identifier used to identify each local segment can be reused within the IGP area since the local segment is only known to that specific router.

Global segments are the opposite of a local segment. Global segments are set domain-wide, it is a part of the segment routing global block for that domain. Such instruction set within a global segment applies to all network nodes within the applied domain. Such global segments are added to all forwarding information bases on all nodes within the domain. [6]

Interior-Gateway Protocol Segment Identifiers

Link state protocols, as explained earlier, are used to distribute global and local segments throughout the domain, thus making such principle very important. Segment routing supports the use of Intermediate System to Intermediate System (IS-IS) and Open Shortest Path First (OSPF). Such technology enables the expression of any path that being, a singular IGP segment, or multiple IGP segments throughout the segment routing domain. An added requirement for the advertisement of IGP segments requires extensions to link-state IGP protocols such as IS-IS and OSPF. Such extensions will be explained later in this reading. [6]

Interior-Gateway Protocol Segment – Prefix-SID

A Prefix-SID is a segment that includes information referring to the network prefix, algorithm, and topology. A Prefix-SID is global unless explicitly configured otherwise, within a segment routing domain. Each packet that enters the domain with an active Prefix-SID will be forwarded in a way dependent on the constraint-based shortest-path calculation applied. Although such a packet is always forwarded along the ECMP-aware shortest path since both constraint-based shortest path calculations available being 'shortest path first' and 'strict shortest path first' utilize the ECMP-aware SPF algorithm.

The shortest-path-first algorithm noted within the Prefix-SID is the default. Such packet is forwarded using the ECMP-aware shortest path employed by the IGP/s used within the domain. This algorithm although, allows for a node along the path to apply another forwarding decision based on its own local policy.

The strict shortest path-first algorithm forces the packet to be forwarded through the means of the ECMP-aware algorithm, thus ignoring all local forwarding policies applied by nodes on the path. Such a strict algorithm ensures that the path for all packets remains unchanged and not to potentially be altered as according to the previously noted algorithm. [5]

Prefix-SIDs are further divided into the following segment identifiers:

- Anycast-SID: An Anycast-SID is used to forward packets towards the closest node of the stated nodes defined within the anycast set. An Anycast set is referred to a set of routers. Such a segment identifier is useful within traffic engineering due to the simplicity surrounding expressing micro-engineering policies.
- Node-SID: A type of Prefix-SID that refers to a specific node, such as a router. Such an identifier notes the exact prefix of a node's loopback interface and has global significance. [6]

Adjacency-SID

An Adjacency-SID, also known as Adj-SID, is a segment identifier of local significance. Such identifier points to a specific link within the same domain, this being a specific interface and the next-hop out of that interface. Once segment routing is enabled over an interior-gateway protocol for an address-family, for any interface that such IGP traverses through, that specific address-family will automatically allocate an adjacency-SID towards all neighbours out of that specific interface. [5]

Segment Routing Operations

As mentioned briefly, during the explanation of the segment routing process, there are three actions that segment routing capable nodes perform on received segments. Such actions relate closely to the actions performed in MPLS networks, on MPLS labels.

- Continue: Also known as MPLS Swap within SR-MPLS, this action is executed when the active segment is still active, but not completed. Such operations tend to occur within global segments due to such segment, potentially including the use of multiple hops.
- Push: Known as MPLS Push within SR-MPLS. The action 'Push' inserts a segment at the top of the segment stack.
- Next: Known as MPLS Pop within SR-MPLS. Such action is executed when the active segment is completed, and the next segment in the stack is ready for inspection. [6]

Source Routing

The use of source routing is present in how the operation of segment routing is conducted. Source routing is the act of decision making surrounding the forwarding path being made by the source router. The main benefits provided by leveraging the source routing paradigm is that the nodes within the network, in this case, routers, do not have to maintain routing information within their routing table due to the forwarding steps being specified within each segment. [2]

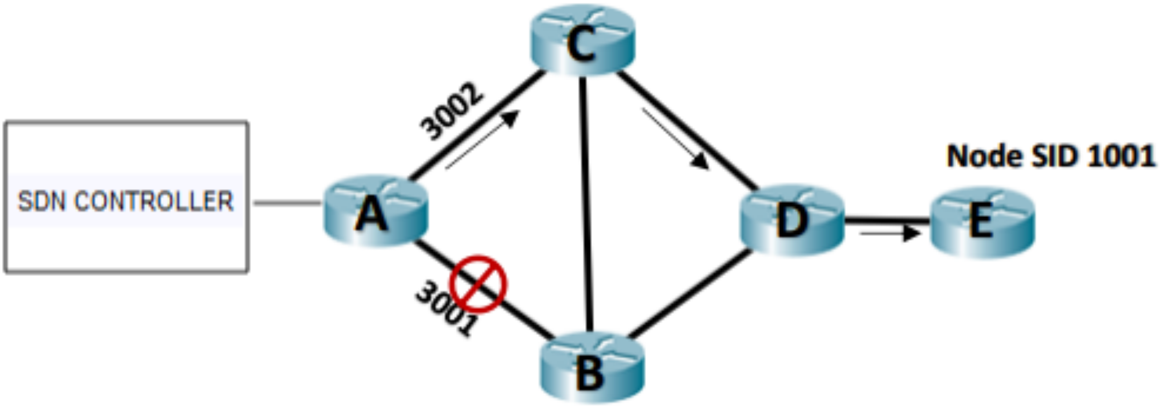
Segment Routing Use Cases

The following section aims to describe use cases surrounding the operation of segment routing in-order to obtain maximum benefits through such operation.

Traffic Engineering – Segment Routing Tunnels

Within segment routing, tunnels can be created to target specific customer needs, such as service-level agreements. Segment routing tunnels can be implemented in a way that provides maximum benefits such as increased throughput and network performance through the use of integrating different SIDs.

When implementing traffic engineering, the use of providing path avoidance is one of the most useful tools. In segment routing, the exploitation of adjacency SIDs allows the network operator to specify a specific path for traffic to flow through the network, thus implementing path avoidance. Such typical implementation of the previously noted is the following – [6]



For this example, a packet would like to be sent to Router E – (Node SID 1001). A typical segment routing implementation will push the node segment to the top of the packet thus it will be forwarded according to the shortest path. Although the use of the Node segment identifier utilizes the shortest path through the means of the ECMP-aware SPF algorithm, as mentioned before. [5] So, in this case, the packet has potentially two paths to take to reach the destination being Router E, and that is RA->RC->RD, and RA->RB->RD. In the situation of a link becoming overloaded, such as the link between RA and RB, the use of path avoidance allows the SDN controller to dynamically push the packet towards Router C to avoid an overloaded link. [6]

Anycast SIDs can also be used to provide traffic engineering. As mentioned before, an anycast SID consists of a list of routers that the packet flows through to reach its destination. Such use of SIDs allows for network environments such as ISPs to express macro engineering policies. Examples of macro engineering policies within dual-plane networks include the likes of “Flow 3 injected in node F toward node E must go via plane 2”. [7]

Fast Reroute

As mentioned previously, segment routing can be made to target a customer’s specific needs, that being a service-level agreement or else. Such service-level agreement may be seen as ‘tight,’ thus to help meet service-level agreement terms, the implementation of fast reroutes is needed. Such implementation is a local protection mechanism in-which provides quick connectivity restoration after a sudden failure of network components. [8]

The implementation of Topology Independent Loop-Free Alternate (TI-LFA) is used to provide fast reroute capabilities within segment routing networks. Such benefits granted by using TI-LFA are the following –

- 100-percent convergence, 50-msec link, SRLG, and node protection
- Automatically computed by IGP
- Prevents congestion
- Protects IP and LDP traffic. [9]

For TI-LFA operation, the protection path is automatically computed by the IGP being utilized within the segment routing network. The protection path utilizes a post-convergence path, which is the next best path to be used in the case of a primary path failure. Post-convergence paths are set by the network operator to support the rerouting of traffic in the situation of a path failure. In TI-LFA, packets are rerouted via attaching backup segments. [9]

MPLS service transport

The implementation of segment routing to an MPLS data plane allows for the use of tunnels to more simply transport services such as VPWs, VPNs, and VPLs, while only utilizing protocols such as IS-IS and OSPF. Thus, noting that to deliver such services, the use of RSVP-TE and LDP is no longer needed. This type of usage surrounding segment routing grants benefits such as providing simplicity within operation through the means of only needing one intra-domain protocol for operation along with not needing to support IGP synchronization extensions. Also, such usage provides improved scaling through the use of a single Node-SID for each node, thus reducing the number of LSDB entries. [8]

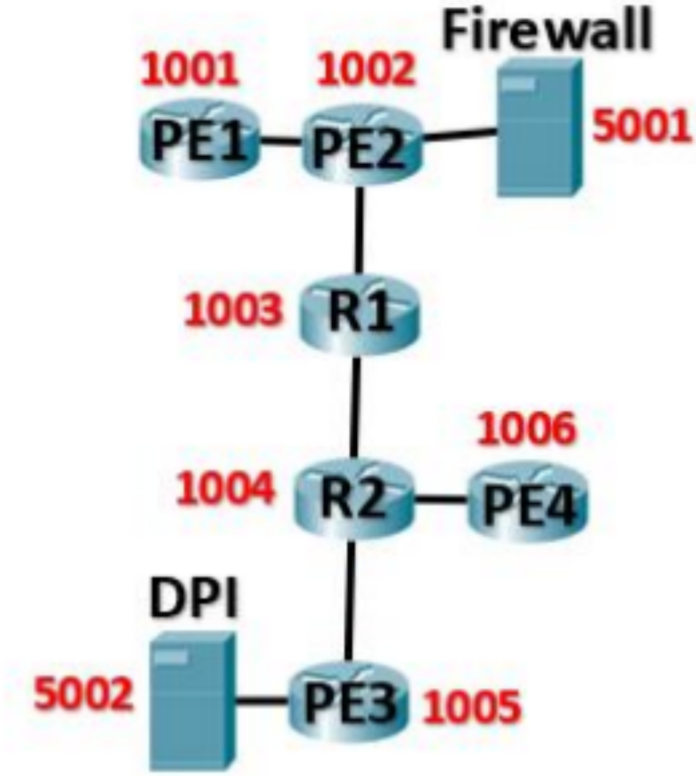
Service chaining

Service chaining can be implemented through the use of segment routing. Such implementation is used to move packets through services offered by other appliances. These services may include accounting, firewalling, etc. These ‘other appliances’ tend to be implemented independently from routers, thus due to the nature of such appliance, they need to be updated, replaced, and migrated frequently. Introducing multiple different appliances within a network may cause a large management overhead for the operators, due to such implementation of appliances being extremely tight, and not supporting the flexibility needed for dynamic behaviors within the network. To face this specific problem, a technique called ‘Service Function Chaining’ was proposed. This technique instantiates a service function path, which is a list of service functions in the packet header, thus removes the constraints surrounding physical topology-based service functions. Segment routing already provides a similar framework to how service chaining operates, so it’s seen as a suitable candidate. [10]

The following is an example of a service chaining operation within a segment routing network.

The operator of the following network would like to apply a set of services, in a fixed order, for traffic between R1 and R2:

- Firewall policies
- Deep packet inspection



When segment routing is utilizing service chaining, the service segments are defined by each SR node directly connected. In this scenario, the service segment for the Firewall is defined as 5001 by PE2, and DPI is defined as 5002 by PE3. These defined service segments have a local significance, it is not global.

In this network, PE1 wants to send a packet to PE4. Thus, the packet has to transverse through R1 and R2. To implement service chaining into such network, PE1 will push a segment header with the following list – 1002, 5001, 1005, 5002, 1006 then send the packet. PE2 will receive the packet then forward it to the service segment denoted by ‘5001’. Once PE2 receives the packet back from the firewall, the packet will be forwarded to PE3, then to the deep packet inspection service denoted by ‘5002’. Once received back from the deep packet inspection service, PE3 will then forward the packet to PE4 – the destination. [10]

Segment Routing Tradeoffs

This section focuses on the design tradeoffs surrounding the use of segment routing within networks. Prior to this section, it has been seen that multiple tradeoffs of prior similar technology have occurred, and a large amount of these tradeoffs are seen as positive.

Firstly, unlike MPLS, when utilizing segment routing, the programming of forwarding information, allocation, and distribution of segments are all carried out by the IGP extensions. This causes a tradeoff surrounding the requirement of heavy protocols such as RSVP-TE and LDP for signaling. This is one of the biggest tradeoffs that provide the most noticeable benefits, the benefits include the following:

- Network Overhead is reduced
- Improvement of fast reroute technologies
- The complexity surrounding the control plane is diminished
- The scalability and functionality of the network are increased. [11]

Touching base with the concept of MPLS again, the implementation of tunnels is used, which is one of the main causes surrounding congestion as MPLS networks grow. Segment routing can be seen as a new concept that utilizes MPLS without the use of additional tunnels. Instead, the implementation of source routing is used to calculate temporary tunnels from the source to the destination, which remain active as long as traffic is traveling through them. Which again, grants a positive outcome by avoiding one of the main causes of congestion within an MPLS-only network. [10]

MPLS label stacks are removed and replaced with a close competitor being segment lists, that operate similarly. As explained previously, segment lists contain a list of segment identifiers, where the active segment is the segment at the top of the list. [10]

Verdict

With the introduction of MPLS in the 90s, we saw a powerful tunnelling mechanism come into play, with such tunnel functionality being the fundamental success of delivering support for services such as MPLS-based VPNs. Although such introduction came with many drawbacks, which lead to the creation of segment routing. As noted previously, segment routing aims to improve upon MPLS by encoding explicit path information into packets at the ingress router, thus not warranting the need to store additional information on routers along the path. Such concept resolved the major issue of scalability within MPLS-TE. Segment routing provides a scalable and flexible architecture aimed at supporting constant evolving network requirements. As presented throughout this paper, the concept surrounding MPLS was given, highlighting the main drawbacks presented which lead to the creation of segment routing. An introduction surrounding segment routing was given, with how such technology operates within today's networks, to provide an easy comparison between both segment routing and MPLS architectures. Technological principles presented within segment routing were discussed, highlighting the differences surrounding how segment routing works when compared to MPLS, such as the use of segment identifiers, source routing, and local/global significant segments. Cases in-which how segment routing can be implemented into networks were presented, thus noting more benefits that segment routing accomplishes within networks such as the use of service chaining, fast reroutes using TI-LFA, and the improvement of network throughput, and performance within traffic engineering. Lastly the trade-offs between the movement from a pure-MPLS network to a segment routing network were noted, noting that such 'trade-offs' were seen to only provide a benefit within the network.

Bibliography

[1] Cisco, “Introduction to Cisco Segment Routing Tutorial,” 2016.

[2] G. Maila, “Segment Routing,” IEEE, 2017.

[3] E. Moreno, “Traffic engineering in segment routing networks,” 2017.

[4] ISP Workshops, “Comparing IS-IS and OSPF,” 2016.

[5] Internet Engineering Task Force, “Segment Routing Architecture,” 2019.

[6] A. Kos, “SEGMENT ROUTING PRINCIPLES AND APPLICATIONS FOR SDN,” 2015.

[7] Cisco, “Introduction to Segment Routing,” 2019.

[8] Network Working Group, “Segment Routing Use Cases,” 2014.

[9] K. M. -. CISCO, “Demo - Topology Independent - Loop Free Alternate (TI-LFA),” 2016.

[10] C. Filsfils, “The Segment Routing Architecture,” 2015.

[11] Sreejith, “What is segment routing? What are its benefits and applications?,” 5 October