Network Working Group
Internet-Draft
Intended status: Informational
Expires: 12 September 2026

C. Schmutzer, Ed.
Z. Ali, Ed.
Cisco Systems, Inc.
P. Maheshwari
Airtel India
R. Rokui
Ciena
A. Stone
Nokia

11 March 2026

Circuit Style Segment Routing Policy
draft-ietf-spring-cs-sr-policy-17

Abstract

This document describes how Segment Routing (SR) policies can be used to satisfy the requirements for bandwidth, end-to-end recovery and persistent paths within a SR network. The association of two co-routed unidirectional SR Policies satisfying these requirements is called "Circuit Style" SR Policy (CS-SR Policy).

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

IP services typically leverage ECMP and local protection. However, packet transport services (commonly referred to as "private lines") that are delivered via pseudowires such as [RFC4448], [RFC4553], [RFC9801], [RFC5086] and [RFC4842] for example, require:

* Persistent end-to-end bidirectional traffic engineered paths that provide predictable and near-symmetric latency

* A requested amount of bandwidth per path that is assured irrespective of changing network utilization from other services

* Fast end-to-end protection and restoration mechanisms

* Monitoring and maintenance of path integrity

* Data plane remaining up while control plane is down

Such a "transport centric" behavior is referred to as "Circuit Style" in this document.

This document describes how Segment Routing (SR) Policies [RFC9256] and adjacency segment identifiers (adjacency-SIDs) defined in the SR architecture [RFC8402] together with a centralized controller such as a stateful Path Computation Element (PCE) [RFC8231] can be used to satisfy those requirements. It includes how end-to-end recovery and path integrity monitoring can be implemented.

A Circuit Style SR Policy (CS-SR Policy) is an association of two co-routed unidirectional SR Policies satisfying the above requirements and allowing for a SR network to carry both typical IP (connection-less) services and connection-oriented transport services.

2. Requirements Notation

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

3. Terminology

* BSID : Binding Segment Identifier

* CS-SR : Circuit Style Segment Routing

* DWDM : Dense Wavelength Division Multiplexing

* ID : Identifier

* LSP : Label Switched Path

* LSPA : LSP Attributes

* NRP : Network Resource Partition

* OAM : Operations, Administration and Maintenance

* OF : Objective Function

* PCE : Path Computation Element

* PCEP : Path Computation Element Communication Protocol

* PT : Protection Type

* SID : Segment Identifier

* SLA : Service Level Agreement

* SDH : Synchronous Digital Hierarchy

* SONET : Synchronous Optical Network

* SR : Segment Routing

* STAMP : Simple Two-Way Active Measurement Protocol

* TI-LFA : Topology Independent Loop Free Alternate

* TLV : Type Length Value

4. Reference Model

The reference model for CS-SR Policies follows the SR architecture [RFC8402] and SR Policy architecture [RFC9256] and is depicted in Figure 1.

```
                         +----------------+
         +-------------->|   controller   |<------------+
         |               +----------------+             |
   PCEP/BGP/config                               PCEP/BGP/config
         |                                              |
         v   <<<<<<<<<<<<<< CS-SR Policy >>>>>>>>>>>>>  v
   +-------+                                          +-------+
   |       |=========================================>|       |
   |   A   | SR Policy from A to Z                    |   Z   |
   |       |<=========================================|       |
   +-------+                    SR Policy from Z to A +-------+
```

Figure 1: Circuit Style SR Policy Reference Model

Given the nature of CS-SR Policies, paths are computed and maintained by a centralized entity providing a consistent simple mechanism for initializing the co-routed bidirectional end-to-end paths, performing bandwidth allocation control, as well as monitoring facilities to ensure SLA compliance for the life of the CS-SR Policy.

CS-SR Policies can be instantiated in the headend routers by using PCEP or BGP as a communication protocol between the headend routers and the central controller or by configuration.

* When using PCEP as the communication protocol, the controller is a stateful PCE as defined in [RFC8231] and SR policy candidate paths are signaled using the PCEP extensions defined in [RFC9862]. When using SR-MPLS [RFC8660], PCEP extensions defined in [RFC8664] are used. When using SRv6 [RFC8754] [RFC8986], PCEP extensions defined in [RFC9603] are used.

* When using BGP as the communication protocol, the BGP extensions defined in [RFC9830] are used.

* When using configuration, an appropriate YANG model such as [I-D.ietf-spring-sr-policy-yang] can be used.

To satisfy the requirements of CS-SR Policies, each link in the topology used by or intended to support CS-SR Policies MUST have:

* An adjacency-SID which is:

   - Persistent, which could be statically configured or auto-generated: to ensure that its value does not change after an event that may cause dynamic states to change (e.g. router reboot).

   - Non-protected: to avoid any local TI-LFA protection [RFC9855] to happen upon interface/link failures.

* The bandwidth available for CS-SR Policies specified.

* A per-hop behavior ([RFC3246] or [RFC2597]) that ensures that the specified bandwidth is always available to CS-SR Policies independent of any other traffic.

To ensure deterministic traffic placement onto parallel physical links and Operations, Administration, and Maintenance (OAM) per physical link, an dedicated adjacency-SID SHOULD be assigned to each physical link.

This means when using link bundles (i.e. [IEEE802.1AX]), a adjacency-SID is assigned per L2 member-link using the mechanisms described in [RFC8668] and [RFC9356]. And that parallel L3 adjacencies described in Section 3.4.1 of [RFC8402] are not used.

This is not needed when the traffic carried by a CS-SR Policy has enough entropy ([RFC6391], [RFC6790], [RFC6437]) for traffic load-balancing across multiple member-links to work well.

When using SR-MPLS [RFC8660], existing IGP extensions defined in [RFC8667] and [RFC8665] and BGP-LS defined in [RFC9085] can be used to distribute the topology information including those persistent and unprotected adjacency-SIDs.

When using SRv6 [RFC8754], the IGP extensions defined in [RFC9352] and [RFC9513] and BGP-LS extensions in [RFC9514] apply.

4.1. Managing Bandwidth

In a network, resources are represented by links of certain bandwidth. In a circuit switched network such as Synchronous Optical Network (SONET) / Synchronous Digital Hierarchy (SDH), Optical Transport Network (OTN) or Dense Wave Division Multiplexing (DWDM) resources (timeslots or a wavelength) are allocated for a provisioned connection at the time of reservation even if no communication is present. In a packet switched network, resources are only allocated when communication is present, i.e. packets are to be sent. This allows for the total reservations to exceed the link bandwidth and can in general lead to link congestion and packet loss.

To satisfy the bandwidth requirement for CS-SR Policies it MUST be ensured that packets carried by CS-SR Policies can always be sent up to the reserved bandwidth on each hop along the path.

This is done by:

* Firstly, CS-SR Policy bandwidth reservations per link MUST be limited to equal or less than the physical link bandwidth. For time-scheduled (TS) reservations ([RFC8413]) this has to be true for a given time window.

* Secondly, ensuring traffic for each CS-SR Policy is limited to the bandwidth reserved for that CS-SR Policy by traffic policing or shaping and admission control on the ingress of the pseudowire.

* Thirdly, ensuring that during times of link congestion CS-SR Policy traffic is not buffered or dropped.

For the third step several approaches can be considered:

* Allocate a dedicated physical link of bandwidth P to CS-SR Policies and allow CS-SR reservations up to bandwidth C. Consider bandwidth N allocated for network control, ensure that P - N >= C.

* Allocate a dedicate logical link (i.e. 801.q VLAN on ethernet) to CS-SR Policies on a physical link of bandwidth P. Limit the total utilization across all other logical links to bandwidth O by traffic policing or shaping and ensure that P - N - O >= C.

* Allocate a dedicated Diffserv codepoint [RFC2597] to map traffic of CS-SR Policies into a specific queue not used by any other traffic.

* Use of dedicated persistent unprotected adjacency-SIDs that are solely used by CS-SR traffic, managed by network design and policy (which is outside the scope of this document). These dedicated SIDs used by CS-SR Policies MUST NOT be used by features such as TI-LFA [RFC9855] for defining the repair path and microloop avoidance for defining the loop-free path.

For networks with low CS-SR traffic volume the approach of a dedicated physical link is undesirable and the option of using a dedicated logical link or dedicated Diffserv codepoint is preferred. If the number of L3 adjacencies in the network is a concern the use of a dedicated Diffserv codepoint is preferred over the use of a dedicated logical link.

The approach of allocating a Diffserv codepoint can leverage any of the following Per-Hop Behavior (PHB) strategies below, where P is the bandwidth of a physical link, N is the bandwidth allocated for network control and C is the bandwidth reserved for CS-SR policies:

* Use a Assured Forwarding (AF) class queue [RFC2597] for CS-SR Policies and limit the total utilization across all other queues to bandwidth O by traffic policing or shaping and ensure that P - N - O >= C.

* Use a Expedited Forwarding (EF) class queue [RFC3246] for CS-SR Policies and limit the total utilization across all other EF queues of higher or equal priority to bandwidth O by traffic policing or shaping and ensure that P - N - O >= C.

* Use a Expedited Forwarding (EF) class queue for CS-SR Policies with a priority higher than all other EF queues and limit the utilization of the CS-SR Policy EF queue by traffic policing to C <= P - N.

The use of a dedicated Diffserv codepoint for CS-SR traffic requires the marking of all traffic steered into CS-SR Policies on the ingress with that specific codepoint consistently across the domain.

In addition, the headends MAY measure the actual bandwidth utilization of a CS-SR Policy to raise alarms when bandwidth utilization thresholds are passed or to request the reserved bandwidth to be adjusted. Using telemetry collection the alarms or bandwidth adjustments can also be triggered by the controller.

Additional background information on general traffic engineering principles can be found in [RFC9522].

5. CS-SR Policy Characteristics

A CS-SR Policy has the following characteristics:

* Requested bandwidth: bandwidth to be reserved for the CS-SR Policy

  - Bandwidth may be adjusted after initial creation as long as no change in path is required

  - Multiple segment-lists may be instantiated to satisfy the bandwidth requirement

* Bidirectional co-routed: a CS-SR Policy between headends A and Z is an association of an SR Policy from A to Z and an SR Policy from Z to A following the same path(s)

* Deterministic and persistent paths: segment lists with strict hops using unprotected adjacency-SIDs that can be statically configured or auto-generated.

* Not automatically recomputed or reoptimized: the segment list of a candidate path MUST NOT change automatically to a segment list representing a different path (for example upon topology change).

* More than one candidate paths in case of protection/restoration:

  - Following the SR Policy architecture, the highest preference valid path is carrying traffic.

  - Depending on the protection/restoration scheme (Section 9), lower priority candidate paths

    o may be pre-computed.

    o may be pre-programmed.

    o may need to be disjoint.

  - Protection switching, restoration and reversion behavior is bidirectional

* It is RECOMMENDED that candidate paths only contain one segment list to avoid asymmetrical routing due to independent load balancing across multiple segment lists on each headend.

* Continuity check and performance measurement are activated on each candidate path (Section 10) and performed per segment-list.

6. CS-SR Policy Creation

6.1. Policy Creation when using PCEP

6.1.1. PCC-initiated Mode

Considering the scenario illustrated in Figure 1 a CS-SR Policy between headends A and Z is instantiated by configured a SR Policy on both headend A (with Z as endpoint) and headend Z (with A as endpoint).

Both headend routers A and Z act as PCC and delegate path computation to the PCE using PCEP with the procedures described in Section 5.7.1 of [RFC8231] and [RFC9862]. For SR-MPLS the extensions defined in [RFC8664] are used. And SRv6 specific extensions are defined in [RFC9603].

The functional requirements of an CS-SR Policy expressed in Section 5 are signaled using PCEP extensions defined in [RFC5440], [RFC8800], [I-D.ietf-pce-sr-bidir-path], [I-D.ietf-pce-circuit-style-pcep-extensions] and [I-D.ietf-pce-multipath].

The candidate paths of the CS-SR Policy are reported and updated following PCEP procedures of [RFC8231].

6.1.2. PCE-initiated Mode

The CS-SR Policy can be instantiated in the network between headends A and Z by a PCE using PCE-initiated procedures defined in [RFC8281]. For PCE-initiated procedures no SR Policy configuration is required on the headends A and Z acting as PCC.

The PCE performs path computation in line with the functional requirements expressed in Section 5 and requests the headends A and Z to initiate a SR Policy using the PCEP extensions listed in Section 6.1.1.

Following initiation, the candidate paths of the CS-SR Policy are reported and updated following PCEP procedures of [RFC8231] and share the same behavior as the PCC-initiated mode.

Connectivity verification and performance measurement is enabled via local policy configuration on the headends, as there is no standard signaling mechanism available.

6.2. Policy Creation when using BGP

Considering the scenario illustrated in Figure 1, instead of configuring SR Policies on both headend A (with Z as endpoint) and headend Z (with A as endpoint), a CS-SR Policy between A and Z is instantiated by a request (e.g. application API call) to the controller.

The controller performs path computation in line with the functional requirements expressed in Section 5 and instantiates the SR Policies in headends A and Z using the BGP extensions defined in [RFC9830].

Connectivity verification and performance measurement is enabled via local policy configuration on the headends, as there is no standard signaling mechanism available.

6.3. Maximum SID Depth Constraint

The segment lists used by CS-SR Policy candidate paths are constrained by the maximum number of segments a router can impose onto a packet.

When using SR-MPLS this constraint is called "Base MPLS Imposition MSD" and is advertised via IS-IS [RFC8491], OSPF [RFC8476], BGP-LS [RFC8814] and PCEP [RFC8664].

When using SRv6 this constraint is called "SRH Max H.encaps MSD" and is advertised via IS-IS [RFC9352], OSPF [RFC9513], BGP-LS [RFC9514] and PCEP [RFC9603].

The MSD constraint is typically resolved by leveraging a segment list reduction technique, such as using Node SIDs and/or Binding SIDs (BSIDs) (SR architecture [RFC8402]) in a segment list, which represents one or many hops in a given path.

As described in Section 5, adjacency-SIDs without local protection are used in CS-SR Policies to ensure that there is no per-hop ECMP, no localized rerouting due to topological changes, and no invocation of localized protection mechanisms, as the alternate path may not be providing the desired SLA.

If a CS-SR Policy path requires segment list reduction, a SR Policy can be programmed in a transit node, and its BSID can be used in the segment list of the CS-SR Policy, if the following requirements are met:

* The transit SR Policy is unprotected, hence only has one candidate path.

* The transit SR Policy follows the rerouting and optimization characteristics defined in Section 5 which implies the segment list of the candidate path MUST only use unprotected adjacency-SIDs.

This ensures that traffic for CS-SR Policies using a BSID does not get locally rerouted due to topological changes or locally protected due to failures. A transit SR Policy may be pre-programmed in the network or automatically injected in the network by a PCE.

7. CS-SR Policy State Reporting

CS-SR Policy state reporting by the headend routers back to the central controller is essential to confirm success or failure of the instantiation and making the controller aware of any state changes throughout the lifetime of the CS-SR Policy in the network.

The headend routers can report CS-SR Policy state by using

* PCEP procedures of [RFC8231].

* BGP-LS procedures of [RFC9857].

* an appropriate YANG model such as [I-D.ietf-spring-sr-policy-yang].

8. CS-SR Policy Deletion

8.1. Policy Deletion when using PCEP

When using PCC-initiated mode, the headends A and Z send a PCRpt message with the R flag set to 1 to inform the PCE about the deletion of a candidate path.

When using PCE-initiated mode, the PCE does send a PCInitiate message to the headends A and Z and to instruct them to delete a candidate path.

8.2. Policy Deletion when using BGP

The controller withdraws a candidate path per [RFC9830] to instruct headends A and Z to delete a candidate path.

9. Recovery Schemes

Various recovery (protection and restoration) schemes can be implemented for a CS-SR Policy. As described in Section 4.3 of [RFC4427], there is a subtle distinction between the terms "protection" and "restoration" based on the resource allocation done during the recovery path establishment. The same definitions apply for CS-SR Policy recovery schemes, wherein:

* Protection: another candidate path is computed and fully established in the data plane and ready to carry traffic.

* Restoration: a candidate path may be computed and may be partially established but is not ready to carry traffic.

The term "failure" is used to represent both "hard failures" such complete loss of connectivity detected by continuity check described in Section 10.1 or degradation, i.e., when the packet loss ratio increased beyond a configured acceptable threshold.

For candidate path establishment the procedures described in Section 6, for candidate path tear down the procedures in Section 8 and for state reporting the procedures in Section 7 can be used.

9.1. Unprotected

In the most basic scenario, no protection or restoration is required. The CS-SR Policy has only one candidate path.

In case of a failure along the path the CS-SR Policy will go down and traffic will not be recovered.

Typically, two CS-SR Policies are deployed either within the same network with disjoint paths or in two separate networks and the overlay service is responsible for traffic recovery.

As soon as the failure(s) that brought the candidate path down are cleared, the candidate path is activated, traffic is sent across it and state is reported accordingly.

9.2. 1:1 Protection

For fast recovery against failures the CS-SR Policy has two candidate paths. Both paths are established but only the candidate with higher preference is activated and is carrying traffic. The second candidate path MUST be computed disjoint to the first candidate path and programmed as backup in the forwarding plane as described in Section 9.3 of [RFC9256].

Upon a failure impacting the candidate path with higher preference carrying traffic, the candidate path with lower preference is activated immediately and traffic is now sent across it.

Protection switching is bidirectional. As described in Section 10.1, both headends will generate and receive their own loopback mode test packets, hence even a unidirectional failure will always be detected by both headends without protection switch coordination required.

Two cases are to be considered when the failure condition impacting a candidate path with higher preference has cleared:

* Revertive switching: automatically re-activate the higher preference candidate path after a configurable period of time and start sending traffic over it.

* Non-revertive switching: do not activate the higher preference candidate path and keep sending traffic via the lower preference candidate path unless manually requested by the operator.

9.3. Restoration

9.3.1. 1+R Restoration

Similarly to 1:1 protection described in Section 9.2, in this recovery scheme the CS-SR Policy has two candidate paths.

To avoid pre-allocating protection bandwidth by the controller ahead of failures, but still being able to recover traffic flow over an alternate path through the network in a deterministic way (maintaining the required bandwidth commitment), the second candidate path with lower preference is established "on demand" and activated upon failure of the first candidate path.

As bandwidth reservations for failed candidate paths are not freed, resource allocation in the network can be optimized, by the second candidate path sharing bandwidth reservations with the first candidate path on links that were not affected by the failures.

As soon as failure(s) that brought the first candidate path down are cleared, the second candidate path is getting torn down and traffic is reverted to the first candidate path.

Restoration and reversion behavior is bidirectional. As described in Section 10.1, both headends use continuity check in loopback mode and therefore, even in case of unidirectional failures, both headends will detect the failure or clearance of the failure and switch traffic away from the failed or to the recovered candidate path.

9.3.2. 1:1+R Restoration

For further resiliency in case of multiple concurrent failures when using 1:1 protection described in Section 9.2 that could bring down both candidate paths, a third candidate path (in this section referred to as "R") with a preference lower than the other two candidate paths (in this section referred to as first and second candidate path) is added to the CS-SR Policy to enable restoration for double failure cases.

There are two possible operating models:

* R established upon double failure

  As in Section 9.3.1, to avoid pre-allocating any additional bandwidth by the controller ahead of double failures, the third candidate path may only be requested when both candidate paths are affected by failures.

  As soon as either the first or second candidate path recovers, traffic will be reverted and the third candidate path MUST be torn down.

* R pre-established after single failure

  Alternatively, the third candidate path can also be requested, pre-computed and programmed as backup already whenever either the first or second candidate path go down with the downside of more bandwidth being set aside ahead of time. When doing so, the third candidate path MUST be computed disjoint to the still operational candidate path.

  The third candidate path will get activated and carry traffic when further failures lead to both the first and second candidate path being down.

  As long as either the first or the second candidate path is active, the third candidate path is kept, updated (if needed) to ensure diversity to the active candidate path and is not carrying traffic.

  Once both, the first and the second candidate path have recovered, the third candidate path is torn down.

As noted in Section 9.3.1, resource allocation in the network can be optimized, by the third candidate path sharing bandwidth reservations with the failed candidate paths on links that were not affected by the failures.

Again, restoration and reversion behavior is bidirectional. As described in Section 10.1, both headends use continuity check in loopback mode and therefore even in case of unidirectional failures both headends will detect the failure or clearance of the failure and switch traffic away from the failed or to the recovered candidate path.

10. Operations, Administration, and Maintenance (OAM)

10.1. Continuity Check

The continuity check for each segment list on both headends MAY be done using

* Simple Two-Way Active Measurement Protocol (STAMP) in loopback measurement mode as described in section 6 and the session state described in section 11 of [I-D.ietf-spring-stamp-srpm-mpls] for SR-MPLS and [I-D.ietf-spring-stamp-srpm-srv6] for SRv6.

* Bidirectional Forwarding Detection (BFD) [RFC5880].

* Seamless BFD (S-BFD) [RFC7880].

The use of STAMP is RECOMMENDED as it leverages a single protocol for both continuity check and performance measurement (see Section 10.2 of this document) and allows for a single session to be used, depending on the desired performance measurement session mode (two-way described in section 4, one-way described in section 5 or loopback described in section 6 of [I-D.ietf-spring-stamp-srpm-mpls] for SR-MPLS and [I-D.ietf-spring-stamp-srpm-srv6] for SRv6).

As the STAMP test packets are including both the segment list of the forward and reverse path, standard segment routing data plane operations will make those packets get forwarded along the forward path to the tailend and along the reverse path back to the headend.

To be able to send STAMP test packets for loopback measurement mode, the STAMP Session-Sender (i.e., the headend) needs to acquire the segment list information of the reverse path:

* When using PCEP, the headend forms the bidirectional SR Policy association using the procedure described in [I-D.ietf-pce-sr-bidir-path] and receives the information about the reverse segment list from the PCE as described in section 4.5 of [I-D.ietf-pce-multipath]

* When using BGP, the controller does inform the headend routers about the reverse segment list using the Reverse Segment List Sub-TLV defined in section 4.1 of [I-D.ietf-idr-sr-policy-path-segment].

For cases where multiple segment lists are used by a candidate path, the headends will declare a candidate path down after continuity check has failed for one or more segment lists because the bandwidth requirement of the candidate path can no longer be met.

10.2. Performance Measurement

Assuming a single STAMP session in loopback mode is used for continuity check and performance measurement, the round-trip delay can be measured and the round-trip loss can be estimated as described in section 8 of [I-D.ietf-spring-stamp-srpm-mpls] for SR-MPLS and [I-D.ietf-spring-stamp-srpm-srv6] for SRv6.

Considering that candidate paths are co-routed, the delay in the forward and reverse direction can be assumed to be similar. Under this assumption, one-way delay can be derived by dividing the round-trip delay by two.

10.3. Candidate Path Validity Verification

A stateful PCE/controller is in sync with the headend routers in the network topology and the CS-SR Policies provisioned on them. As described in Section 5 a path MUST NOT be automatically recomputed by the controller after or optimized for topology changes unless it is a restoration path.

However, there may be a requirement for the stateful PCE/controller to tear down a path if the path no longer satisfies the original requirements, such as insufficient bandwidth, diversity constraint no longer met or latency constraint exceeded and only the stateful PCE/controller can detect this and not the headend routers themselves.

For a CS-SR Policy configured with multiple candidate paths, a headend may switch to another candidate path if the stateful PCE/controller decided to tear down the active candidate path.

11. Operational Considerations

As a Circuit Style SR Policy (CS-SR Policy) is an association of two co-routed unidirectional SR Policies, the manageability considerations outlined in Section 11 of [RFC9256] do apply.

Additional operational considerations are:

* Configure both sides identical (behavior and flags)

* When using PCEP, configure Association ID, Association Source, optional Global Association Source TLV, and optional Extended Association ID TLV according to [RFC8697].

* LSP ping and traceroute [[RFC9716]] is performed unidirectionally (per SR Policy).

* Diversity among candidate paths can be verified by using LSP traceroute.

* CS-SR Policies will lead to more alarms in the fault management system, because a candidate path can stay down until a network topology failure which caused the down event clears.

Configuration and operation can use the YANG model defined in [I-D.ietf-spring-sr-policy-yang].

Further this document is informational as it does not introduce any new mechanism, but rather describes how to use existing mechanisms to create the Circuit Style SR policy. As such the whole document can be considered as a operational guideline.

11.1. External Commands

External commands are typically issued by an operator to control the candidate path state of a CS-SR Policy using the management interface of:

* Headends: When the CS-SR Policy was instantiated via configuration or PCEP PCC-initiated mode

* PCE/controller: When the CS-SR Policy was instantiated via BGP or PCEP PCE-initiated mode

11.1.1. Candidate Path Switchover

Typically used in conjunction with non-revertive protection switching to re-activate a recovered candidate path upon operator request.

It also allows operators to trigger a switch between candidate paths even if no failure is present, e.g., to proactively drain a resource for maintenance purposes.

A operator triggered switching request between candidate paths on a headend is unidirectional and SHOULD be requested on both headends to ensure co-routing of traffic.

11.1.2. Candidate Path Re-computation

While no automatic re-optimization or pre-computation of CS-SR Policy candidate paths is allowed as specified in Section 5, network operators trying to optimize network utilization may explicitly request a candidate path to be re-computed at a certain point in time.

12. Security Considerations

This document does provide guidance on how to implement a CS-SR Policy leveraging existing mechanisms and protocol extensions. As such, it does not introduce any new security considerations.

The MPLS or SRv6 network is assumed to be a trusted and secure domain. Attackers who manage to send spoofed packets into the domain could easily disrupt services leveraging CS-SR Policies. The protections against such attacks are described by considerations in Section 4.2 of [RFC5920] and in Section 8 of [RFC8402].

Security considerations for the SR Policy Architecture defined in Section 10 of [RFC9256] do apply to this document as well.

To satisfy the bandwidth requirement of CS-SR Policies, the Differentiated Service architecture [RFC2475] is leveraged and the security considerations in Section 6 of [RFC2475] do apply. If a dedicated Diffserv codepoint is assigned to CS-SR Policies, the use by any other traffic is to be prevented to ensure QoS is properly enforced.

Further a misconfiguration of requested bandwidth for CS-SR Policies can lead to blocking out other CS-SR Policies from consuming available bandwidth and bandwidth starvation of non-CS-SR traffic.

Depending on how a CS-SR Policy is instantiated and reported, the following security considerations do apply

* PCEP:
  - Section 7 of [RFC8664]
  - Section 6 of [RFC9603]
  - Section 8 of [RFC9862]
  - Section 6 of [I-D.ietf-pce-sr-bidir-path]
  - Section 7 of [I-D.ietf-pce-circuit-style-pcep-extensions]
  - Section 10 of [I-D.ietf-pce-multipath]
  - Section 8 of [I-D.ietf-idr-sr-policy-path-segment]
* BGP:
  - Section 7 of [RFC9830]

   - Section 9 of [RFC9857]

* Configuration:

   - Section 8 of [I-D.ietf-spring-sr-policy-yang]

Depending on the protocol used for OAM, the following security considerations do apply

* STAMP: Section 15 of [I-D.ietf-spring-stamp-srpm-mpls] and [I-D.ietf-spring-stamp-srpm-srv6]

* BFD/S-BFD: Section 9 of [RFC5880] and Section 11 of [RFC7880]

13. IANA Considerations

This document has no IANA actions.

14. Acknowledgements

The author’s want to thank Samuel Sidor, Mike Koldychev, Rakesh Gandhi, Alexander Vainshtein, Tarek Saad, Ketan Talaulikar and Yao Liu for providing their review comments, Yao Liu for her very detailed shepherd review and all contributors for their inputs and support.

15. References

15.1. Normative References

[I-D.ietf-idr-sr-policy-path-segment]
Li, C., Li, Z., Yin, Y., Cheng, W., and K. Talaulikar, "SR Policy Extensions for Path Segment and Bidirectional Path", Work in Progress, Internet-Draft, draft-ietf-idr-sr-policy-path-segment-14, 11 September 2025, <https://datatracker.ietf.org/doc/html/draft-ietf-idr-sr-policy-path-segment-14>.

[I-D.ietf-pce-circuit-style-pcep-extensions]
Sidor, S., Maheshwari, P., Stone, A., Jalil, L., and S. Peng, "Path Computation Element Communication Protocol (PCEP) extensions for Circuit Style Policies", Work in Progress, Internet-Draft, draft-ietf-pce-circuit-style-pcep-extensions-15, 6 March 2026, <https://datatracker.ietf.org/doc/html/draft-ietf-pce-circuit-style-pcep-extensions-15>.

[I-D.ietf-pce-multipath]
Koldychev, M., Sivabalan, S., Saad, T., Beeram, V. P., Bidgoli, H., Peng, S., and S. Sidor, "Path Computation Element Communication Protocol (PCEP) Extensions for Signaling Multipath Information", Work in Progress, Internet-Draft, draft-ietf-pce-multipath-20, 2 March 2026, <https://datatracker.ietf.org/doc/html/draft-ietf-pce-multipath-20>.

[I-D.ietf-pce-sr-bidir-path]
Li, C., Chen, M., Cheng, W., Gandhi, R., and Q. Xiong, "Path Computation Element Communication Protocol (PCEP) Extensions for Associated Bidirectional Segment Routing (SR) LSPs", Work in Progress, Internet-Draft, draft-ietf-pce-sr-bidir-path-25, 6 March 2026, <https://datatracker.ietf.org/doc/html/draft-ietf-pce-sr-bidir-path-25>.

[I-D.ietf-spring-sr-policy-yang]
Saleh, T., Raza, S. K., Zhuang, S., Matsushima, S., and V. P. Beeram, "YANG Data Model for Segment Routing Policy", Work in Progress, Internet-Draft, draft-ietf-spring-sr-policy-yang-06, 20 October 2025, <https://datatracker.ietf.org/doc/html/draft-ietf-spring-sr-policy-yang-06>.

[I-D.ietf-spring-stamp-srpm-mpls]
Gandhi, R., Filsfils, C., Janssens, B., Chen, M., and R. F. Foote, "Performance Measurement Using Simple Two-Way Active Measurement Protocol (STAMP) for Segment Routing over the MPLS Data Plane", Work in Progress, Internet-Draft, draft-ietf-spring-stamp-srpm-mpls-00, 2 October 2025, <https://datatracker.ietf.org/doc/html/draft-ietf-spring-stamp-srpm-mpls-00>.

[I-D.ietf-spring-stamp-srpm-srv6]
Gandhi, R., Filsfils, C., Janssens, B., Chen, M., and R. F. Foote, "Performance Measurement Using Simple Two-Way Active Measurement Protocol (STAMP) for Segment Routing over the IPv6 (SRv6) Data Plane", Work in Progress, Internet-Draft, draft-ietf-spring-stamp-srpm-srv6-00, 2 October 2025, <https://datatracker.ietf.org/doc/html/draft-ietf-spring-stamp-srpm-srv6-00>.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <https://www.rfc-editor.org/rfc/rfc2119>.

[RFC2475] Blake, S., Black, D., Carlson, M., Davies, E., Wang, Z., and W. Weiss, "An Architecture for Differentiated Services", RFC 2475, DOI 10.17487/RFC2475, December 1998, <https://www.rfc-editor.org/rfc/rfc2475>.

[RFC5440] Vasseur, JP., Ed. and JL. Le Roux, Ed., "Path Computation Element (PCE) Communication Protocol (PCEP)", RFC 5440, DOI 10.17487/RFC5440, March 2009, <https://www.rfc-editor.org/rfc/rfc5440>.

[RFC5920] Fang, L., Ed., "Security Framework for MPLS and GMPLS Networks", RFC 5920, DOI 10.17487/RFC5920, July 2010, <https://www.rfc-editor.org/rfc/rfc5920>.

[RFC6391] Bryant, S., Ed., Filsfils, C., Drafz, U., Kompella, V., Regan, J., and S. Amante, "Flow-Aware Transport of Pseudowires over an MPLS Packet Switched Network", RFC 6391, DOI 10.17487/RFC6391, November 2011, <https://www.rfc-editor.org/rfc/rfc6391>.

[RFC6437] Amante, S., Carpenter, B., Jiang, S., and J. Rajahalme, "IPv6 Flow Label Specification", RFC 6437, DOI 10.17487/RFC6437, November 2011, <https://www.rfc-editor.org/rfc/rfc6437>.

[RFC6790] Kompella, K., Drake, J., Amante, S., Henderickx, W., and L. Yong, "The Use of Entropy Labels in MPLS Forwarding", RFC 6790, DOI 10.17487/RFC6790, November 2012, <https://www.rfc-editor.org/rfc/rfc6790>.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <https://www.rfc-editor.org/rfc/rfc8174>.

[RFC8231] Crabbe, E., Minei, I., Medved, J., and R. Varga, "Path Computation Element Communication Protocol (PCEP) Extensions for Stateful PCE", RFC 8231, DOI 10.17487/RFC8231, September 2017, <https://www.rfc-editor.org/rfc/rfc8231>.

[RFC8281] Crabbe, E., Minei, I., Sivabalan, S., and R. Varga, "Path Computation Element Communication Protocol (PCEP) Extensions for PCE-Initiated LSP Setup in a Stateful PCE Model", RFC 8281, DOI 10.17487/RFC8281, December 2017, <https://www.rfc-editor.org/rfc/rfc8281>.

[RFC8402] Filsfils, C., Ed., Previdi, S., Ed., Ginsberg, L., Decraene, B., Litkowski, S., and R. Shakir, "Segment Routing Architecture", RFC 8402, DOI 10.17487/RFC8402, July 2018, <https://www.rfc-editor.org/rfc/rfc8402>.

[RFC8660] Bashandy, A., Ed., Filsfils, C., Ed., Previdi, S., Decraene, B., Litkowski, S., and R. Shakir, "Segment Routing with the MPLS Data Plane", RFC 8660, DOI 10.17487/RFC8660, December 2019, <https://www.rfc-editor.org/rfc/rfc8660>.

[RFC8664] Sivabalan, S., Filsfils, C., Tantsura, J., Henderickx, W., and J. Hardwick, "Path Computation Element Communication Protocol (PCEP) Extensions for Segment Routing", RFC 8664, DOI 10.17487/RFC8664, December 2019, <https://www.rfc-editor.org/rfc/rfc8664>.

[RFC8697] Minei, I., Crabbe, E., Sivabalan, S., Ananthakrishnan, H., Dhody, D., and Y. Tanaka, "Path Computation Element Communication Protocol (PCEP) Extensions for Establishing Relationships between Sets of Label Switched Paths (LSPs)", RFC 8697, DOI 10.17487/RFC8697, January 2020, <https://www.rfc-editor.org/rfc/rfc8697>.

[RFC8754] Filsfils, C., Ed., Dukes, D., Ed., Previdi, S., Leddy, J., Matsushima, S., and D. Voyer, "IPv6 Segment Routing Header (SRH)", RFC 8754, DOI 10.17487/RFC8754, March 2020, <https://www.rfc-editor.org/rfc/rfc8754>.

[RFC8800] Litkowski, S., Sivabalan, S., Barth, C., and M. Negi, "Path Computation Element Communication Protocol (PCEP) Extension for Label Switched Path (LSP) Diversity Constraint Signaling", RFC 8800, DOI 10.17487/RFC8800, July 2020, <https://www.rfc-editor.org/rfc/rfc8800>.

[RFC9256] Filsfils, C., Talaulikar, K., Ed., Voyer, D., Bogdanov, A., and P. Mattes, "Segment Routing Policy Architecture", RFC 9256, DOI 10.17487/RFC9256, July 2022, <https://www.rfc-editor.org/rfc/rfc9256>.

[RFC9603] Li, C., Ed., Kaladharan, P., Sivabalan, S., Koldychev, M., and Y. Zhu, "Path Computation Element Communication Protocol (PCEP) Extensions for IPv6 Segment Routing", RFC 9603, DOI 10.17487/RFC9603, July 2024, <https://www.rfc-editor.org/rfc/rfc9603>.

[RFC9716] Hegde, S., Arora, K., Srivastava, M., Ninan, S., and N. Kumar, "Mechanisms for MPLS Ping and Traceroute Procedures in Inter-Domain Segment Routing Networks", RFC 9716, DOI 10.17487/RFC9716, February 2025, <https://www.rfc-editor.org/rfc/rfc9716>.

[RFC9830] Previdi, S., Filsfils, C., Talaulikar, K., Ed., Mattes, P., and D. Jain, "Advertising Segment Routing Policies in BGP", RFC 9830, DOI 10.17487/RFC9830, September 2025, <https://www.rfc-editor.org/rfc/rfc9830>.

[RFC9857] Previdi, S., Talaulikar, K., Ed., Dong, J., Gredler, H., and J. Tantsura, "Advertisement of Segment Routing Policies Using BGP - Link State", RFC 9857, DOI 10.17487/RFC9857, October 2025, <https://www.rfc-editor.org/rfc/rfc9857>.

[RFC9862] Koldychev, M., Sivabalan, S., Sidor, S., Barth, C., Peng, S., and H. Bidgoli, "Path Computation Element Communication Protocol (PCEP) Extensions for Segment Routing (SR) Policy Candidate Paths", RFC 9862, DOI 10.17487/RFC9862, October 2025, <https://www.rfc-editor.org/rfc/rfc9862>.

15.2. Informative References

[IEEE802.1AX] IEEE, "IEEE Standard for Ethernet", May 2020, <https://ieeexplore.ieee.org/document/9105034>.

[RFC2597] Heinanen, J., Baker, F., Weiss, W., and J. Wroclawski, "Assured Forwarding PHB Group", RFC 2597, DOI 10.17487/RFC2597, June 1999, <https://www.rfc-editor.org/rfc/rfc2597>.

[RFC3246] Davie, B., Charny, A., Bennet, J.C.R., Benson, K., Le Boudec, J.Y., Courtney, W., Davari, S., Firoiu, V., and D. Stiliadis, "An Expedited Forwarding PHB (Per-Hop Behavior)", RFC 3246, DOI 10.17487/RFC3246, March 2002, <https://www.rfc-editor.org/rfc/rfc3246>.

[RFC4427] Mannie, E., Ed. and D. Papadimitriou, Ed., "Recovery (Protection and Restoration) Terminology for Generalized Multi-Protocol Label Switching (GMPLS)", RFC 4427, DOI 10.17487/RFC4427, March 2006, <https://www.rfc-editor.org/rfc/rfc4427>.

[RFC4448] Martini, L., Ed., Rosen, E., El-Aawar, N., and G. Heron, "Encapsulation Methods for Transport of Ethernet over MPLS Networks", RFC 4448, DOI 10.17487/RFC4448, April 2006, <https://www.rfc-editor.org/rfc/rfc4448>.

[RFC4553] Vainshtein, A., Ed. and YJ. Stein, Ed., "Structure-Agnostic Time Division Multiplexing (TDM) over Packet (SAToP)", RFC 4553, DOI 10.17487/RFC4553, June 2006, <https://www.rfc-editor.org/rfc/rfc4553>.

[RFC4842] Malis, A., Pate, P., Cohen, R., Ed., and D. Zelig, "Synchronous Optical Network/Synchronous Digital Hierarchy (SONET/SDH) Circuit Emulation over Packet (CEP)", RFC 4842, DOI 10.17487/RFC4842, April 2007, <https://www.rfc-editor.org/rfc/rfc4842>.

[RFC5086] Vainshtein, A., Ed., Sasson, I., Metz, E., Frost, T., and P. Pate, "Structure-Aware Time Division Multiplexed (TDM) Circuit Emulation Service over Packet Switched Network (CESoPSN)", RFC 5086, DOI 10.17487/RFC5086, December 2007, <https://www.rfc-editor.org/rfc/rfc5086>.

[RFC5880] Katz, D. and D. Ward, "Bidirectional Forwarding Detection (BFD)", RFC 5880, DOI 10.17487/RFC5880, June 2010, <https://www.rfc-editor.org/rfc/rfc5880>.

[RFC7880] Pignataro, C., Ward, D., Akiya, N., Bhatia, M., and S. Pallagatti, "Seamless Bidirectional Forwarding Detection (S-BFD)", RFC 7880, DOI 10.17487/RFC7880, July 2016, <https://www.rfc-editor.org/rfc/rfc7880>.

[RFC8413] Zhuang, Y., Wu, Q., Chen, H., and A. Farrel, "Framework for Scheduled Use of Resources", RFC 8413, DOI 10.17487/RFC8413, July 2018, <https://www.rfc-editor.org/rfc/rfc8413>.

[RFC8476] Tantsura, J., Chunduri, U., Aldrin, S., and P. Psenak, "Signaling Maximum SID Depth (MSD) Using OSPF", RFC 8476, DOI 10.17487/RFC8476, December 2018, <https://www.rfc-editor.org/rfc/rfc8476>.

[RFC8491] Tantsura, J., Chunduri, U., Aldrin, S., and L. Ginsberg, "Signaling Maximum SID Depth (MSD) Using IS-IS", RFC 8491, DOI 10.17487/RFC8491, November 2018, <https://www.rfc-editor.org/rfc/rfc8491>.

[RFC8665] Psenak, P., Ed., Previdi, S., Ed., Filsfils, C., Gredler, H., Shakir, R., Henderickx, W., and J. Tantsura, "OSPF Extensions for Segment Routing", RFC 8665, DOI 10.17487/RFC8665, December 2019, <https://www.rfc-editor.org/rfc/rfc8665>.

[RFC8667] Previdi, S., Ed., Ginsberg, L., Ed., Filsfils, C., Bashandy, A., Gredler, H., and B. Decraene, "IS-IS Extensions for Segment Routing", RFC 8667, DOI 10.17487/RFC8667, December 2019, <https://www.rfc-editor.org/rfc/rfc8667>.

[RFC8668] Ginsberg, L., Ed., Bashandy, A., Filsfils, C., Nanduri, M., and E. Aries, "Advertising Layer 2 Bundle Member Link Attributes in IS-IS", RFC 8668, DOI 10.17487/RFC8668, December 2019, <https://www.rfc-editor.org/rfc/rfc8668>.

[RFC8814] Tantsura, J., Chunduri, U., Talaulikar, K., Mirsky, G., and N. Triantafillis, "Signaling Maximum SID Depth (MSD) Using the Border Gateway Protocol - Link State", RFC 8814, DOI 10.17487/RFC8814, August 2020, <https://www.rfc-editor.org/rfc/rfc8814>.

[RFC8986] Filsfils, C., Ed., Camarillo, P., Ed., Leddy, J., Voyer, D., Matsushima, S., and Z. Li, "Segment Routing over IPv6 (SRv6) Network Programming", RFC 8986, DOI 10.17487/RFC8986, February 2021, <https://www.rfc-editor.org/rfc/rfc8986>.

[RFC9085] Previdi, S., Talaulikar, K., Ed., Filsfils, C., Gredler, H., and M. Chen, "Border Gateway Protocol - Link State (BGP-LS) Extensions for Segment Routing", RFC 9085, DOI 10.17487/RFC9085, August 2021, <https://www.rfc-editor.org/rfc/rfc9085>.

[RFC9352] Psenak, P., Ed., Filsfils, C., Bashandy, A., Decraene, B., and Z. Hu, "IS-IS Extensions to Support Segment Routing over the IPv6 Data Plane", RFC 9352, DOI 10.17487/RFC9352, February 2023, <https://www.rfc-editor.org/rfc/rfc9352>.

[RFC9356] Talaulikar, K., Ed. and P. Psenak, "Advertising Layer 2 Bundle Member Link Attributes in OSPF", RFC 9356, DOI 10.17487/RFC9356, January 2023, <https://www.rfc-editor.org/rfc/rfc9356>.

[RFC9513] Li, Z., Hu, Z., Talaulikar, K., Ed., and P. Psenak, "OSPFv3 Extensions for Segment Routing over IPv6 (SRv6)", RFC 9513, DOI 10.17487/RFC9513, December 2023, <https://www.rfc-editor.org/rfc/rfc9513>.

[RFC9514] Dawra, G., Filsfils, C., Talaulikar, K., Ed., Chen, M., Bernier, D., and B. Decraene, "Border Gateway Protocol - Link State (BGP-LS) Extensions for Segment Routing over IPv6 (SRv6)", RFC 9514, DOI 10.17487/RFC9514, December 2023, <https://www.rfc-editor.org/rfc/rfc9514>.

[RFC9522] Farrel, A., Ed., "Overview and Principles of Internet Traffic Engineering", RFC 9522, DOI 10.17487/RFC9522, January 2024, <https://www.rfc-editor.org/rfc/rfc9522>.

[RFC9801] Gringeri, S., Whittaker, J., Leymann, N., Schmutzer, C., Ed., and C. Brown, "Private Line Emulation over Packet Switched Networks", RFC 9801, DOI 10.17487/RFC9801, July 2025, <https://www.rfc-editor.org/rfc/rfc9801>.

[RFC9855] Bashandy, A., Litkowski, S., Filsfils, C., Francois, P., Decraene, B., and D. Voyer, "Topology Independent Fast Reroute Using Segment Routing", RFC 9855, DOI 10.17487/RFC9855, October 2025, <https://www.rfc-editor.org/rfc/rfc9855>.

Contributors

Daniel Voyer
Bell Canada
Email: daniel.voyer@bell.ca

Luay Jalil
Verizon
Email: luay.jalil@verizon.com

Shuping Peng
Huawei Technologies
Email: pengshuping@huawei.com

Clarence Filsfils
Cisco Systems, Inc.
Email: cfilsfil@cisco.com

Francois Clad
Cisco Systems, Inc.
Email: fclad@cisco.com

Tarek Saad
Cisco Systems, Inc.
Email: tsaad.net@gmail.com

Brent Foster
Cisco Systems, Inc.
Email: brfoster@cisco.com

Bertrand Duvivier
Cisco Systems, Inc.
Email: bduvivie@cisco.com

Stephane Litkowski
Cisco Systems, Inc.
Email: slitkows@cisco.com

Jie Dong
Huawei Technologies
Email: jie.dong@huawei.com

Authors' Addresses

Christian Schmutzer (editor)
Cisco Systems, Inc.
Email: cschmutz@cisco.com

Zafar Ali (editor)
Cisco Systems, Inc.
Email: zali@cisco.com

Praveen Maheshwari
Airtel India
Email: Praveen.Maheshwari@airtel.com

Reza Rokui
Ciena
Email: rrokui@ciena.com

Andrew Stone
Nokia
Email: andrew.stone@nokia.com