SPRING Working Group
Internet Draft
Intended status: Standards Track
Expires: 30 December 2025

Y. Liu
China Mobile
C. Lin
New H3C Technologies
S. Peng
Huawei Technologies
R. Chen
ZTE Corporation
G. Mishra
Verizon Inc.
Y. Qiu
New H3C Technologies

30 June 2025

Flexible Candidate Path Selection of SR Policy
draft-liu-spring-sr-policy-flexible-path-selection-10

Abstract

This document proposes a flexible SR policy candidate path selection method. Based on the real-time resource usage and forwarding quality of candidate paths, the head node can perform dynamic path switching among multiple candidate paths in the SR policy.

Table of Contents

1. Introduction...2
2. Terminology..3
3. Background of requirements...................................3
4. Flexible Candidate Path Selection Method.....................5
   4.1. Threshold Parameters of Candidate Paths.................6
   4.2. Rules for Setting eligibility attribute.................8
   4.3. Flexible Candidate Path Selection Process...............8
5. Usecases of Flexible Candidate Path Selection................9
   5.1. Select the Best Path Based on End-to-End Delay..........9
   5.2. Select the Best Path Based on Available Bandwidth.......10
   5.3. Select the Best Path Based on Actual Bandwidth..........11
6. IANA Considerations..11
7. Security Considerations......................................11
8. References...12
   8.1. Normative References....................................12
   8.2. Informative References..................................12
9. Acknowledgments..14
Authors' Addresses..14

1. Introduction

Segment routing (SR) [RFC8402] is a source routing paradigm that explicitly indicates the forwarding path for packets at the ingress node. The ingress node steers packets into a specific path according to the Segment Routing Policy (SR Policy) as defined in [RFC9256].

An SR Policy may have multiple candidate paths that are provisioned or signaled [I-D.ietf-idr-sr-policy-safi] [RFC8664] from one of more sources. The tie-breaking rules defined in [RFC9256] result in determination of a single "active path" in a formal definition.

Refer to [RFC9256] only the active candidate path MUST be used for forwarding traffic that is being steered onto that policy except for certain scenarios such as fast reroute where a backup candidate path may be used. A candidate path can be represented as a segment list or a set of segment lists. If a set of segment lists is associated with the active path of the policy, then the steering is per flow and weighted-ECMP (W-ECMP) based according to the relative weight of each valid segment list.

According to the criteria for the validity of candidate paths described in Section 5 of [RFC9256], if there is a valid segment list in the active candidate path, the active candidate path is still valid. When some segment lists of the active candidate path are invalid, the active candidate path may still be valid, but it may not continue to meet the actual forwarding requirements.

[I-D.karboubi-spring-sr-policy-eligibility] introduces a concept of an eligibility attribute at the candidate path level, not only at the time of the computation but also through topology and network changes to ensure that user intentions are preserved while carrying service traffic.

This document proposes a method for setting the eligibility attribute, which can influence the selection of candidate paths. For specific preference rules, refer to [I-D.karboubi-spring-sr-policy-eligibility].

Based on real-time resource usage and forwarding quality of candidate paths, the head node can dynamically adjust the eligibility attribute value, enabling it to perform dynamic path switching among multiple candidate paths within the SR policy.

[RFC2386] provides valuable background on QoS-based routing, details some issues and requirements associated with QoS-based routing, and proposes a framework for employing QoS-based routing within the Internet. This document describes an SR Policy mechanism where the path state is switched based on the resource status of the traversed path. However, it does not address the challenges related to dynamic distributed scheduling or resource reservation along intermediate paths. The document specifies the capability to switch to alternative paths within a strategy when the current path fails to satisfy designated link quality criteria, such as bandwidth, delay, or packet loss. In instances where a controller issues an SR Policy encompassing multiple paths, should a path’s link quality not meet the established requirements, a switch to a backup path for forwarding is executed.

2. Terminology

The definitions of the basic terms are identical to those found in Segment Routing Policy Architecture [RFC9256].

3. Background of requirements

When some segment lists of the active candidate path are invalid, according to [RFC9256], if there is a valid segment list in the active candidate path, the active candidate path is still valid. But the paths of remaining segment lists may not meet the SR policy forwarding performance requirements, such as insufficient path bandwidth. Even if there are other candidate paths with lower preference that can meet the forwarding performance requirements in the SR policy, the traffic will continue to be forwarded along the original active candidate path.

As an example, consider the following SR Policy to illustrate the issues present in the current candidate path selection process in detail.

```
SR Policy POL1
   Candidate Path CP1
      Preference 200
      Segment List 1 <SID11...SID1i>, Weight 1
      Segment List 2 <SID21...SID2j>, Weight 1
      Segment List 3 <SID31...SID3k>, Weight 1
   Candidate Path CP2
    Preference 100
      Segment List 4 <SID41...SID4i>, Weight 1
      Segment List 5 <SID51...SID5j>, Weight 1
      Segment List 6 <SID61...SID6k>, Weight 1
```

There are two static candidate paths CP1 and CP2 in SR policy POL1. CP1 has a higher preference. Both candidate paths are composed of three static segment lists with the same weight. The path indicated by each segment list can carry traffic of 100Mbps bandwidth. When all Segment Lists in CP1 are valid, the effective bandwidth of the candidate path is 300Mbps.

The bandwidth of the actual traffic forwarded by the SR policy is between 100Mbps and 150Mbps. Because the traffic forwarded on the candidate path will share the load on the three segment list paths according to the weight value. Therefore, normally, the candidate path can meet the forwarding requirements. The traffic is forwarded on the three segment lists of the high preference candidate paths of the SR policy.

When the segment list 1 and 2 in the high-preference candidate path CP1 are invalid, according to the candidate path validity criteria described in [RFC9256] Section 5, because the segment list 3 in CP1 is still valid, the active candidate path CP1 is still valid. All traffic of SR policy POL1 will continue to be forwarded through the path of CP1. However, because segment list 3 can only forward 100Mbps traffic, over-bandwidth traffic will be discarded.

Of course, when the Segment List path fault is detected, the network device can report the detected fault information to the controller. The controller optimizes the forwarding path after receiving the message. However, this interaction process is relatively long, and it is difficult to meet the requirement for fast switching.

When the quality of high-preference candidate paths deteriorates, due to issues such as insufficient available bandwidth, increased end-to-end transmission delay, or segment lists that fail to meet service requirements, the same need arises. The goal is to switch traffic to other candidate paths within the SR policy that better satisfy the forwarding quality requirements.

To address this issue, this document proposes a new candidate path selection rule that defines resource thresholds and forwarding quality requirements for candidate paths.

If a candidate path does not satisfy the forwarding quality requirements, its eligibility attribute MUST be set to false. During the active CP selection process, the head-end SHALL use this eligibility attribute as an additional mandatory criterion, in conjunction with the rules defined in [RFC9256], Section 2.9. When a CP’s eligibility attribute is false, it indicates that the path cannot forward traffic and therefore MUST NOT be considered for active CP selection.

4. Flexible Candidate Path Selection Method

As described in [RFC9256], the candidate path selection process operates primarily on the candidate path Preference. A candidate path is selected when it is valid and it has the highest Preference value among all the valid candidate paths of the SR Policy.

[I-D.karboubi-spring-sr-policy-eligibility] introduce a new attribute at the candidate path level called eligibility. Only candidate paths with eligibility as true must be considered as part of the active candidate path selection defined in [RFC9256].

This document proposes using forwarding quality requirements and resource requirements of candidate paths as eligibility criteria for path selection.

A headend MAY be informed about the forwarding quality requirements of a candidate path for an SR Policy <Color, Endpoint> through various means, including configuration, PCEP, or BGP. The extensions of BGP and PCEP are described in

[I-D.liu-idr-bgp-sr-policy-cp-threshold] and

[I-D.liu-pce-sr-policy-cp-threshold].

When a candidate path fails to meet forwarding quality requirements, its eligibility attribute should be set to false, thereby excluding it from active candidate path selection.

For candidate paths containing multiple segment lists:

- If a segment list fails to meet forwarding quality requirements, it should be excluded from forwarding operations.

- When all segment lists under a candidate path fail to meet forwarding quality requirements, the path’s eligibility attribute should be set to false, subsequently disqualifying it from active candidate path selection.

4.1. Threshold Parameters of Candidate Paths

The threshold parameters of candidate paths can include but are not limited to the following:

* Jitter

* Latency

* Packet loss

  Delay, jitter, and packet loss are thresholds at the segment list level.

  When the jitter, delay, or packet loss of a valid segment list does not meet the specified threshold requirements, the segment list will be deemed invalid and will no longer participate in

  load sharing traffic.

* Available bandwidth

  The bandwidth threshold is the threshold at the candidate path level.

  CP available bandwidth = CP preset bandwidth * (Sum of Segment List weights in Up state / Sum of all Segment List weights)

* Actual bandwidth

  The actual bandwidth refers to the sum of the actual available remaining bandwidth of each valid segment list in the candidate path.

  Due to the different congestion conditions of each node on the forwarding path, the actual bandwidth that can forward service packets may differ from the preset bandwidth. By utilizing some measurement mechanisms, the actual minimum available bandwidth

  and actual minimum remaining bandwidth of all nodes along the

  path can be obtained. The specific measurement mechanism is not within the scope of this document.

* Precision Availability Metrics (PAM)

Consider a candidate path of SR policy as a Service Level Objective (SLO), based on the Precision Availability Metrics

(PAM) defined in [RFC9544], determine whether the candidate path meets the forwarding requirements.

If both segment list-level thresholds (such as latency, jitter, or packet loss) and candidate path-level thresholds (such as available bandwidth) are specified, then when one or more segment lists in the candidate path fail to meet the segment list-level thresholds, it indicates that these segment lists cannot provide forwarding capabilities that meet the SLA requirements. These segment lists will be marked as unavailable and will no longer participate in packet forwarding. After excluding these segment lists, it should be verified whether the candidate path still meets the forwarding quality requirements. If it does, traffic can continue to be forwarded along that candidate path.

For example, two threshold parameters, delay and available bandwidth, are specified simultaneously for the candidate path with multiple segment lists. When the delay of a segment list exceeds the threshold, the following processing is performed:

1) Remove the segment list from the forwarding path first.

2) Calculate the current available bandwidth of CP based on the weight ratio of the remaining effective segment lists and the bandwidth of CP.

3) Check whether the current available bandwidth of CP still meets the bandwidth threshold requirements.

   * If the available bandwidth still meets the requirements, the candidate path still meets the forwarding quality

     requirements, and the traffic is still forwarded along this candidate path.

   * Otherwise, set the eligibility attribute of this CP to false. The system should then consider switching service traffic to another active candidate path with better forwarding quality.

If the candidate path does not specify any threshold parameters, select the primary candidate path according to the selection method defined in [RFC9256].

By default, there is no threshold parameter specified on the candidate path.

4.2. Rules for Setting eligibility attribute

When a candidate path's current forwarding quality meets specified threshold requirements, its eligibility attribute MUST be set to true, indicating this path is valid for:

* Traffic forwarding operations.

* Active candidate path selection (per [RFC9256] selection methodology)

Conversely, when a candidate path fails to meet quality requirements, its eligibility attribute MUST be set to false.

For candidate paths without defined threshold parameters:

* The eligibility attribute MUST default to true.

* Primary path selection follows [RFC9256] procedures.

When multiple eligible candidate paths coexist in an SR policy:

* Only paths with eligibility=true MAY participate in active

  path selection.

* Detailed behavior specified in

  [I-D.karboubi-spring-sr-policy-eligibility].

4.3. Flexible Candidate Path Selection Process

The process of selecting the best candidate path for SR policy through the threshold parameter of the candidate path is as follows.

1) Configure the threshold parameters on the candidate path of the head node through static manual configuration or controller distribution.

2) The head node monitors whether the available resources and forwarding quality of the SR policy candidate path exceed the thresholds.

3) The forwarding quality of path can be obtained through active or passive performance measurement methods, such as iOAM [RFC9378], STAMP [I-D.ietf-spring-stamp-srpm], TWAMP [RFC5375], etc. The real-time quality data can be calculated by the controller and distributed to the head node, or calculated by the head node according to the network measurement data. The measurement method and quality data acquisition method are beyond the scope of this document.

4) According to the rules described in Section 4.2, when the available resources are less than the threshold, or the forwarding quality cannot meet the threshold requirements, select a new active candidate path.

5) After the old active candidate path eliminates the fault or improves the forwarding quality, whether to recover can be specified by the configuration. If fault recovery is required, start a wait timer for delay recovery. When the timer expires and the old active candidate path still meets the threshold requirements, the traffic will be switched to the old higher preference candidate path.

For avoiding path switching frequently, both over-threshold switching and fault recovery should be delayed. The interval of delay waiting can be adjusted by configuration.

5. Usecases of Flexible Candidate Path Selection

The SR policy in Section 3 is still used to illustrate how the flexible candidate path selection method switches candidate paths.

SR policy POL1 has two candidate paths CP1 and CP2. The Preference of CP1 is 200, and the Preference of CP2 is 100. Both candidate paths are composed of three segment lists with the same weight.

5.1. Select the Best Path Based on End-to-End Delay

The quality requirement for the services carried on the SR policy is that the transmission delay must be less than 200ms. The bandwidth of the actual traffic forwarded by the SR policy is between 100Mbps and 150Mbps.

When the delay of Segment List 1 does not meet the requirements, continue to check the available bandwidth of CP1. Due to segment list 2 only having 100Mbps bandwidth, it cannot meet the actual traffic forwarding requirements. CP1’s eligibility attribute MUST be set to false, triggering the selection of CP2 as POL1’s new active candidate path. The traffic forwarded by POL1 is switched to the path of CP2 for forwarding.

```
SR Policy POL1
  Candidate Path CP1
     Preference 200
     Delay threshold 200ms //Delay<=200ms
     Segment List 1 <SID11...SID1i>, Weight 1 //100M, Delay>1s
     Segment List 2 <SID21...SID2i>, Weight 1 //100M, Delay<100ms
  Candidate Path CP2
     Preference 100
     Delay threshold 200ms  //Delay<=200ms
     Segment List 3 <SID31...SID3i>, Weight 1 //100M, Delay<100ms
```

```
           Segment List 4 <SID41...SID4i>, Weight 1 //100M, Delay<100ms
```

5.2. Select the Best Path Based on Available Bandwidth

The preset bandwidth for CP1 and CP2 is both 300Mbps. Each segment list can carry a maximum of 100Mbps traffic. The quality requirement for service traffic is that the available bandwidth of the forwarding path must not be less than 150Mbps.

```
SR Policy POL1
   Candidate Path CP1
      Preference 200
      Preset bandwidth 300Mbps
      Available bandwidth threshold 150Mbps
      Segment List 1 <SID11...SID1i>, Weight 1
      Segment List 2 <SID21...SID2j>, Weight 1
      Segment List 3 <SID31...SID3k>, Weight 1
   Candidate Path CP2
    Preference 100
      Preset bandwidth 300Mbps
    Available bandwidth threshold 150Mbps
      Segment List 4 <SID41...SID4i>, Weight 1
      Segment List 5 <SID51...SID5j>, Weight 1
      Segment List 6 <SID61...SID6k>, Weight 1
```

First, take the available bandwidth as the threshold parameter of POL1. The threshold for configuring the available bandwidth is 150Mbps. When the available bandwidth of the candidate path is less than 150Mbps, perform path switching.

Normally, the three segment lists of CP1 and CP2 are valid. The available bandwidth of CP1 is 300Mbps, and the available bandwidth can meet the threshold requirements. CP1’s eligibility attribute MUST be set to true, CP1 is selected as the active candidate path according to the Preference.

If the paths indicated by Segment List 1 and 2 fail, Segment List 1 and 2 become invalid, and the available bandwidth of CP1 becomes 100Mbps. Because the available bandwidth of CP1 is lower than the specified threshold, CP1 has failed to meet the forwarding quality requirements. CP1’s eligibility attribute MUST be set to false. Need to reselect the active candidate path for POL1.

The three segment lists of the low-preference candidate path CP2 of POL1 are valid, and the available bandwidth can meet the threshold requirements. CP2’s eligibility attribute MUST be set to true. CP2 is selected as the new active candidate path of POL1. The traffic forwarded by POL1 will switch to the path of CP2 for forwarding.

5.3. Select the Best Path Based on Actual Bandwidth

In scenarios involving the actual available bandwidth measurement method for SRv6, as described in

[I-D.liu-ippm-srv6-bandwidth-measurement], the quality requirement for the services carried on the SR policy mandates that the actual available bandwidth of the forwarding path must exceed 80 Mbps. If traffic congestion occurs on a node in Segment List 1, resulting in a maximum forwarding capacity of only 50 Mbps for service traffic, and if Segment List 2 is either in a down state or has exceeded the delay threshold, Segment List 2 will not participate in load sharing traffic.

When the aggregate available bandwidth of CP1 falls below 80 Mbps:

* CP1’s eligibility attribute MUST be set to false.

* CP2’s eligibility attribute MUST be set to true (provided it meets forwarding requirements).

* CP2 SHALL become POL1’s new active candidate path.

```
SR Policy POL1
  Candidate Path CP1
     Preference 200
      Preset bandwidth 200Mbps
     Actual available bandwidth threshold 80Mbps
      Segment List 1 <SID11...SID1i>, Weight 1
         (Actual available bandwidth is only 50Mbps.)
      Segment List 2 <SID21...SID2j>, Weight 1
         (In Down state, or the delay has exceeded the threshold.)
  Candidate Path CP2
     Preference 100
     Preset bandwidth 300Mbps
     Actual available bandwidth threshold 80Mbps
      Segment List 3 <SID41...SID4i>, Weight 1 (100Mbps)
      Segment List 4 <SID51...SID5j>, Weight 1 (100Mbps)
      Segment List 5 <SID61...SID6k>, Weight 1 (100Mbps)
```

The traffic forwarded by POL1 will switch to the path of CP2 for forwarding.

6. IANA Considerations

This document has no IANA actions.

7. Security Considerations

This document does not introduce any security considerations.

8. References

8.1. Normative References

[I-D.ietf-idr-sr-policy-safi] Previdi, S., Filsfils, C., Talaulikar, K., Mattes, P., and Jain, D., "Advertising Segment Routing Policies in BGP", draft-ietf-idr-sr-policy-safi-13 (work in progress), February 2025.

[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/info/rfc2119>.

[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/info/rfc8402>.

[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/info/rfc8174>.

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

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

8.2. Informative References

[I-D.liu-idr-bgp-sr-policy-cp-threshold] Liu, Y., Lin, C., Qiu, Y., " BGP Extension for Distributing CP Threshold Constraints of SR Policy", draft-liu-idr-bgp-sr-policy-cp-threshold-02 (work in progress), November 2024.

[I-D.liu-pce-sr-policy-cp-threshold] Liu, Y., Lin, C., Qiu, Y., " PCEP Extension to Support Signaling Candidate Path Threshold Constraints of SR Policy", draft-liu-pce-sr-policy-cp-threshold-03 (work in progress), February 2025.

[I-D.liu-ippm-srv6-bandwidth-measurement] Liu, Y., Lin, C., Qiu, Y., Liu, Y., Liang, Y., " Measurement Method for Bandwidth of SRv6 Forwarding Path", draft-liu-ippm-srv6-bandwidth-measurement (work in progress), November 2024.

[I-D.karboubi-spring-sr-policy-eligibility] Karboubi, A., Shah, H., Sivalaban, S., Stone, A. and Schmutz, C., "Eligibility Concept in Segment Routing Policies", draft-karboubi-spring-sr-policy-eligibility-02 (work in progress), June 2025.

[I-D.ietf-spring-stamp-srpm] 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 Networks", Work in Progress, Internet-Draft, draft-ietf-spring-stamp-srpm-19, 20 June 2025, <https://datatracker.ietf.org/doc/html/draft-ietf-spring-stamp-srpm-19>.

[RFC2386] Crawley, E., Nair, R., Rajagopalan, B. and H. Sandick, "A Framework for QoS-based Routing in the Internet", RFC 2386, August 1998.

[RFC5375] Hedayat, K., Krzanowski, R., Morton, A., Yum, K., Babiarz, J., "A Two-Way Active Measurement Protocol (TWAMP)", RFC 5375, DOI 10.17487/RFC5375, October 2008, <https://www.rfc-editor.org/info/rfc5375>.

[RFC9378] Brockners, F., Bhandari, S., Bernier, D., Mizrahi, T., "In Situ Operations, Administration, and Maintenance (IOAM) Deployment", RFC 9378, DOI 10.17487/RFC9378, April 2023, <https://www.rfc-editor.org/info/rfc9378>.

[RFC9544] Mirsky, G., Halpern, J., Min, X., Clemm, A., Strassner, J., Francois, J., "Precision Availability Metrics for Services Governed by Service Level Objectives (SLOs)", RFC 9544, DOI 10.17487/RFC9544, March 2024, <https://www.rfc-editor.org/info/rfc9544>.

9. Acknowledgments

The authors would like to thank the following for their valuable contributions of this document:

TBD

Authors' Addresses

Yisong Liu
China Mobile
Beijing
China

Email: liuyisong@chinamobile.com

Changwang Lin
New H3C Technologies
Beijing
China

Email: linchangwang.04414@h3c.com

Shuping Peng
Huawei Technologies
Beijing
China
Email: pengshuping@huawei.com

Ran Chen
ZTE Corporation
Nanjing
China
Email: chen.ran@zte.com.cn

Gyan S. Mishra
Verizon Inc.
Email: gyan.s.mishra@verizon.com

Yuanxiang Qiu
New H3C Technologies
Beijing
China

Email: qiuyuanxiang@h3c.com