

Networking Working Group
Internet-Draft
Intended status: Standards Track
Expires: January 3, 2018

L. Ginsberg
P. Psenak
S. Previdi
Cisco Systems
M. Pilka
July 2, 2017

Segment Routing MPLS Conflict Resolution
draft-ietf-spring-conflict-resolution-05.txt

Abstract

In support of Segment Routing (SR) for an MPLS data plane routing protocols advertise a variety of identifiers used to define the segments which direct forwarding of packets. In cases where the information advertised by a given protocol instance is either internally inconsistent or conflicts with advertisements from another protocol instance a means of achieving consistent forwarding behavior in the network is required. This document defines the policies used to resolve these occurrences.

Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC 2119](#) [[RFC2119](#)].

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of [BCP 78](#) and [BCP 79](#).

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on January 3, 2018.

Internet-Draft

sr-conflict-resolution

July 2017

Copyright Notice

Copyright (c) 2017 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to [BCP 78](#) and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1.	Introduction	3
2.	SR Global Block Inconsistency	3
3.	SR-MPLS Segment Identifier Conflicts	5
3.1.	SID Preference	6
3.2.	Conflict Types	7
3.2.1.	Prefix Conflict	7
3.2.2.	SID Conflict	9
3.3.	Preference rule for resolving conflicts	12
3.4.	Conflict Resolution Algorithm	13
3.5.	Example Behavior - Single Topology/Address Family/Algorithm	14
3.6.	Example Behavior - Multiple Topologies	15
3.7.	Guaranteeing Database Consistency	16
3.8.	Minimizing the occurrence of conflicts	16
4.	Scope of SR-MPLS SID Conflicts	16
5.	Conflict Resolution and non-forwarding nodes	17
6.	Security Considerations	17
7.	IANA Consideration	18
8.	Acknowledgements	18
9.	References	18
9.1.	Normative References	18
9.2.	Informational References	19
Appendix A.	Alternative SID Conflict Resolution Policy Discussion	19
A.1.	Policy: Ignore conflicting entries	19
A.2.	Policy: Preference Algorithm/Quarantine	19

A.3.	Policy: Preference algorithm/ignore overlap only	20
A.4.	Example Behavior - Single Topology/Address Family/Algorithm	20
A.5.	Evaluation of Policy Alternatives	21
	Authors' Addresses	22

[1.](#) Introduction

Segment Routing (SR) as defined in [[SR-ARCH](#)] utilizes forwarding instructions called "segments" to direct packets through the network. Depending on the forwarding plane architecture in use, routing protocols advertise various identifiers which define the permissible values which can be used as segments, which values are assigned to specific prefixes, etc. Where segments have global scope it is necessary to have non-conflicting assignments - but given that the advertisements may originate from multiple nodes the possibility exists that advertisements may be received which are either internally inconsistent or conflicting with advertisements originated by other nodes. In such cases it is necessary to have consistent resolution of conflicts network-wide in order to avoid forwarding loops.

This document is limited to discussion of conflict resolution for identifiers used in an MPLS data plane.

The problem to be addressed is protocol independent i.e., segment related advertisements may be originated by multiple nodes using different protocols and yet the conflict resolution MUST be the same on all nodes regardless of the protocol used to transport the advertisements.

The remainder of this document defines conflict resolution policies which meet these requirements. All protocols which support SR MUST adhere to the policies defined in this document.

[2.](#) SR Global Block Inconsistency

In support of an MPLS dataplane [[SR-MPLS](#)] routing protocols advertise an SR Global Block (SRGB) which defines a set of label ranges reserved for use by the advertising node in support of SR. The details of how protocols advertise this information can be found in the protocol specific drafts e.g., [[SR-OSPF](#)], [[SR-OSPFv3](#)], [[SR-IS-](#)

IS], and [SR-BGP]. However the protocol independent semantics are illustrated by the following example:

Ginsberg, et al.

Expires January 3, 2018

[Page 3]

Internet-Draft

sr-conflict-resolution

July 2017

The originating router advertises the following ranges:

Range 1: (100, 199)
Range 2: (1000, 1099)
Range 3: (500, 599)

The receiving routers concatenate the ranges and build the Segment Routing Global Block (SRGB) as follows:

SRGB = (100, 199)
 (1000, 1099)
 (500, 599)

The indices span multiple ranges:

index=0 means label 100
...
index 99 means label 199
index 100 means label 1000
index 199 means label 1099
...
index 200 means label 500
...

Note that the ranges are an ordered set - what labels are mapped to a given index depends on the placement of a given label range in the set of ranges advertised.

For the set of ranges to be usable the ranges MUST be disjoint.

Sender behavior is defined in various SR protocol drafts such as [SR-IS-IS] which specify that senders MUST NOT advertise overlapping ranges.

Receivers of SRGB ranges MUST validate the SRGB ranges advertised by other nodes. If the advertised ranges do not conform to the restrictions defined in the respective protocol specification receivers MUST ignore all advertised SRGB ranges from that node. Operationally the node is treated as though it did not advertise any SRGB ranges. [SR-MPLS] defines the procedures for mapping global SIDs to outgoing labels.

Note that utilization of local SIDs (e.g. adjacency SIDs) advertised by a node is not affected by the state of the advertised SRGB.

3. SR-MPLS Segment Identifier Conflicts

In support of an MPLS dataplane Segment Identifiers (SIDs) are advertised and associated with a given prefix. SIDs may be advertised in the prefix reachability advertisements originated by a routing protocol (PFX) . SIDs may also be advertised by a Segment Routing Mapping Server (SRMS). How this is done is defined in the protocol specific drafts e.g., [SR-OSPF], [SR-OSPFv3], [SR-IS-IS], and [SR-BGP]

Information in a SID advertisement is used to construct a mapping entry. A generalized mapping entry can be represented using the following definitions:

Prf - Preference Value (See [Section 3.1](#))

Pi - Initial prefix

Pe - End prefix

L - Prefix length

Lx - Maximum prefix length (32 for IPv4, 128 for IPv6)

Si - Initial SID value

Se - End SID value

R - Range value (See Note 1)

T - Topology
A - Algorithm (see [[SR-ARCH](#)])

A Mapping Entry is then the tuple: (Prf, Pi/L, Si, R, T, A)
 $Pe = (Pi + ((R-1) \ll (Lx-L))$
 $Se = Si + (R-1)$

NOTE 1: The SID advertised in a prefix reachability advertisement always has an implicit range of 1.

NOTE 2: IPv4/IPv6 addresses can be viewed as 32/128 bit integers. Where operations such as addition, subtraction, and/or bit shifting are specified for prefixes this should be interpreted as operations on the integer representation of a prefix.

Note: Topology is a locally scoped identifier assigned by each router. Although it may have an association with Multitopology Identifiers (MTID) advertised by routing protocols it is NOT equivalent to these identifiers. MTIDs are scoped by a given routing protocol. MTID ranges are protocol specific and there may be standardized protocol specific MTID assignments for topologies of a specific type (e.g., an AFI specific topology). As mapping entries can be sourced from multiple protocols it is not possible to use a

network scoped identifier for a topology when storing mapping entries in the local database.

Conflicts in SID advertisements may occur as a result of misconfiguration. When conflicts occur, it is not possible for routers to know which of the conflicting advertisements is "correct". In order to avoid forwarding loops and/or blackholes, there is a need for all nodes to resolve the conflicts in a consistent manner. This in turn requires that all routers have identical sets of advertisements and that they all use the same selection algorithm. This document defines procedures to achieve these goals.

[3.1.](#) SID Preference

If a node acts as an SRMS, it MAY advertise a preference to be

associated with all SRMS SID advertisements sent by that node. The means of advertising the preference is defined in the protocol specific drafts e.g., [[SR-OSPF](#)], [[SR-OSPFv3](#)], and [[SR-IS-IS](#)]. The preference value is an unsigned 8 bit integer with the following properties:

- 0 - Reserved value indicating advertisements from that node MUST NOT be used.
- 1 - 255 Preference value

Advertisement of a preference value is optional. Nodes which do not advertise a preference value are assigned a preference value of 128.

All SIDs advertised in prefix reachability advertisements originated by an IGP implicitly have a preference value of 192.

All SIDs advertised in prefix reachability advertisements originated by BGP implicitly have a preference value of 64.

These preference values are deliberately chosen to favor SID advertisements originated within a domain (IGP and SRMS) over SID advertisements which may have been imported from other domains (BGP). In addition, as BGP originated advertisements may not be known on all nodes within a domain (because not every node will be a BGP speaker), the presence of a BGP originated mapping entry MUST NOT cause a mapping entry originated within the domain to become unusable as this would introduce inconsistency in the set of SIDs considered usable by a node which has the BGP originated mapping entries and the set considered usable by nodes without the BGP originated mapping entries.

[3.2.](#) Conflict Types

Two types of conflicts may occur - Prefix Conflicts and SID Conflicts. Examples are provided in this section to illustrate these conflict types and generic definitions of algorithms to determine when there is a conflict are presented.

[3.2.1.](#) Prefix Conflict

When different SIDs are assigned to the same prefix we have a "prefix conflict". Prefix conflicts are limited to mapping entries sharing the same topology, algorithm, address-family, and prefix length.

[3.2.1.1](#). Prefix Conflict Examples

The simplest example is when two advertisements with a range of 1 assign different SIDs to the same prefix.

Example PC1

```
(192, 192.0.2.120/32, 200, 1, 0, 0)
(192, 192.0.2.120/32, 30, 1, 0, 0)
```

The prefix 192.0.2.120/32 has been assigned two different SIDs:
200 by the first advertisement
30 by the second advertisement

Example PC2

```
(192, 2001:DB8::1/128, 400, 1, 2, 0)
(192, 2001:DB8::1/128, 50, 1, 2, 0)
```

The prefix 2001:DB8::1/128 has been assigned two different SIDs:
400 by the first advertisement
50 by the second advertisement

Prefix conflicts may also occur as a result of overlapping prefix ranges.

Example PC3

(128, 192.0.2.1/32, 200, 200, 0, 0)
(128, 192.0.2.121/32, 30, 10, 0, 0)

Prefixes 192.0.2.121/32 - 192.0.2.130/32 are assigned two different SIDs:

320 through 329 by the first advertisement
30 through 39 by the second advertisement

Example PC4

(128, 2001:DB8::1/128, 400, 200, 2, 0)
(128, 2001:DB8::121/128, 50, 10, 2, 0)

Prefixes 2001:DB8::121/128 - 2001:DB8::130/128 are assigned two different SIDs:

420 through 429 by the first advertisement
50 through 59 by the second advertisement

Examples PC3 and PC4 illustrate a complication - only part of the range advertised in the first advertisement is in conflict. It is logically possible to consider the sub-range(s) which are in conflict as unusable while considering the sub-range(s) not in conflict as usable.

A variant of the overlapping prefix range is a case where we have overlapping prefix ranges but no actual prefix conflict.

Example PC5

(128, 192.0.2.1/32, 200, 200, 0, 0)
(128, 192.0.2.121/32, 320, 10, 0, 0)

(128, 2001:DB8::1/128, 400, 200, 2, 0)
(128, 2001:DB8::121/128, 520, 10, 2, 0)

Although there is prefix overlap between the two IPv4 entries (and the two IPv6 entries) the same SID is assigned to all of the shared prefixes by the two entries.

[3.2.1.2](#). Prefix Conflict Generic Algorithm

The following generic algorithm can be used to determine when any two mapping entries have Prefix Conflicts and what the set of prefixes in conflict are.

Given two mapping entries:

(Prf, P1/L1, S1, R1, T1, A1) and
(Prf, P2/L2, S2, R2, T2, A2)

where $P1 \leq P2$

a prefix conflict exists if all of the following are true:

1) Topologies, algorithms, and prefix lengths are identical

$(T1 == T2) \ \&\& \ (A1 == A2) \ \&\& \ (L1 == L2)$

2) The prefixes are in the same address-family.

3) If there are overlapping prefixes in the two ranges and
if there are different SIDs assigned to any of the prefixes
in the overlapping range

$(P1e \geq P2) \ \&\& \ ((S1 + ((P2 - P1) \gg (Lx - L1)) \neq S2)$

Prefixes in the following range are in conflict:

P2 through $\text{MIN}(P1e, P2e)$

[3.2.2. SID Conflict](#)

When the same SID has been assigned to multiple prefixes we have a "SID conflict". SID conflicts are independent of address-family, independent of prefix len, independent of topology, and independent of algorithm.

[3.2.2.1. SID Conflict Examples](#)

The simplest example is when two mapping entries with a range of 1 assigns different SIDs to the same prefix.

Internet-Draft

sr-conflict-resolution

July 2017

Example SC1

(192, 192.0.2.1/32, 200, 1, 0, 0)

(192, 192.0.2.222/32, 200, 1, 0, 0)

SID 200 has been assigned to 192.0.2.1/32 by the first advertisement.

The second advertisement assigns SID 200 to 192.0.2.222/32.

Example SC2

(192, 2001:DB8::1/128, 400, 1, 2, 0)

(192, 2001:DB8::222/128, 400, 1, 2, 0)

SID 400 has been assigned to 2001:DB8::1/128 by the first advertisement.

The second advertisement assigns SID 400 to 2001:DB8::222/128

SID conflicts may also occur as a result of overlapping SID ranges.

Example SC3

(128, 192.0.2.1/32, 200, 200, 0, 0)

(128, 198.51.100.1/32, 300, 10, 0, 0)

SIDs 300 - 309 have been assigned to two different prefixes.

The first advertisement assigns these SIDs to 192.0.2.101/32 - 192.0.2.110/32.

The second advertisement assigns these SIDs to 198.51.100.1/32 - 198.51.100.10/32.

Example SC4

(128, 2001:DB8::1/128, 400, 200, 2, 0)

(128, 2001:DB8:1::1/128, 500, 10, 2, 0)

SIDs 500 - 509 have been assigned to two different prefixes.

The first advertisement assigns these SIDs to 2001:DB8::101/128 - 2001:DB8::10A/128.

The second advertisement assigns these SIDs to 2001:DB8:1::1/128 - 2001:DB8:1::A/128.

Examples SC3 and SC4 illustrate a complication - only part of the range advertised in the first advertisement is in conflict.

SID conflicts may also occur because the same SID has been used in two different algorithms, two different topologies, two different address families, or prefixes with two different lengths.

Ginsberg, et al.

Expires January 3, 2018

[Page 10]

Internet-Draft

sr-conflict-resolution

July 2017

Example SC5

(128, 192.0.2.1/32, 200, 1, 0, 0)
(128, 192.0.2.1/32, 200, 1, 0, 1)

SID 200 has been assigned to the same prefix with two different algorithms.

Example SC6

(128, 192.0.2.1/32, 200, 1, 0, 0)
(128, 2001:DB8::1/128, 200, 1, 0, 0)

SID 200 has been assigned to prefixes in two different address-families.

[3.2.2.2](#). SID Conflict Generic Algorithm

The following generic algorithm can be used to determine when any two mapping entries have SID Conflicts and what the set of SIDs in conflict are.

Given two mapping entries:

(Prf, P1/L1, S1, R1, T1, A1) and
(Prf, P2/L2, S2, R2, T2, A2)

a SID conflict exists if all of the following are true:

1) If the SID ranges overlap

(S1 <= S2) && (S1e >= S2)

2) If the same SID is assigned to prefixes with different address-families, prefix lengths, topologies, or algorithms or the same SID is assigned to two different prefixes for any of the prefixes in either range.

P1 and P2 are NOT in the same address family OR
L1 != L2 OR
T1 != T2 OR
A1 != A2 OR
(P1 + ((S1e-S2) << (L1x-L1))) != P2

SIDs in the following range are in conflict:

S2 through MIN(S1e,S2e)

Ginsberg, et al.

Expires January 3, 2018

[Page 11]

Internet-Draft

sr-conflict-resolution

July 2017

[3.3.](#) Preference rule for resolving conflicts

When a conflict is detected the following algorithm is used to select the preferred mapping entry. Evaluation is made in the order specified. Prefix conflicts are evaluated first. SID conflicts are then evaluated on the Active entries remaining after Prefix Conflicts have been resolved.

1. Higher preference value wins
2. Smaller range wins
3. IPv6 entry wins over IPv4 entry
4. Longer prefix length wins
5. Smaller starting address (considered as an unsigned integer value) wins
6. Smaller algorithm wins
7. Smaller starting SID wins
8. If topology IDs are NOT identical both entries MUST be ignored

When applying the preference rule to prefix/SID pairs associated with an advertised mapping entry with a range greater than one, each prefix/SID pair in the range is considered as having the range associated with the advertised mapping entry. For example:

Advertised mapping entry: (128, 192.0.2.1/32, 200, 200, 0, 0)

The advertisement covers 200 prefix/SID pairs:

192.0.2.1/32 200

192.0.2.2/32 201

...

192.0.2.200/32 399

Each of these prefix/SID pairs is considered as having a range of 200 when applying Rule #2 above.

As SIDs associated with prefix reachability advertisements have a preference of 192 and an implied range of 1 while by default SRMS preference is 128, the default behavior is then to prefer SIDs advertised in prefix reachability advertisements over SIDs advertised by SRMSs, but an operator can choose to override this behavior by setting SRMS preference higher than 192.

Preferring advertisements with smaller range has the nice property that a single misconfiguration of an SRMS entry with a large range will not be preferred over a large number of advertisements with smaller ranges.

Since topology identifiers are locally scoped, it is not possible to make a consistent choice network wide when all elements of a mapping entry are identical except for the topology. This is why both entries MUST be ignored in such cases (Rule #8 above). Note that Rule #8 only applies when considering SID conflicts since Prefix conflicts are restricted to a single topology.

[3.4.](#) Conflict Resolution Algorithm

The following logical steps MUST be followed in the order specified when resolving conflicts.

Step 1: Resolve Prefix Conflicts (same topology/address family/

algorithm)

For each supported topology/address family/algorithm examine all qualifying mapping entries in the following order:

- 1)Preference (start w highest)
- 2)Range (start w smallest)
- 3)Prefix length (start w longest)
- 4)Address (start w smallest)
- 5)SID (start w smallest)

At each step if a prefix conflict is detected the losing prefix/SID pair is declared Inactive and is not considered in any subsequent steps. The remaining prefix/SID pairs are Active.

Mapping entries with Active prefix/SID pairs after completion of Step 1 are fed into ...

Step 2: SID Conflicts (across all topologies/address families/ algorithms)

Examine all Active prefix/SID pairs from Step #1 in the following order:

- 1)Preference (start w highest)
- 2)Range (start w smallest)
- 3)IPv6 entries
 - a)Prefix length (start w longest)
 - b)Address (start w smallest)
- 4)IPv4 entries
 - a)Prefix Length (start w longest)
 - b)Address (start w smallest)
- 5)Algorithm (start w smallest)
- 6)SID (start w smallest)

Prefix/SID pairs which are identical and are associated with the same topology are duplicates - both entries MUST be considered as Active.

Prefix/SID pairs which are identical and are associated with different topologies MUST both be considered Inactive.

Active Entries in the database may be used in forwarding. Inactive entries MUST NOT be used in forwarding.

Note that when the database of mapping entries changes the full set of logical steps MUST be reapplied to the entire database as conflict resolution is NOT transitive.

NOTE: Clever implementors may realize optimizations when rerunning the algorithm by evaluating changed entries as to whether they have potential conflicts with any of the existing entries in the database (both active and inactive). Such optimizations are outside the scope of this specification. The normative behavior is defined by the logical algorithm above.

[3.5.](#) Example Behavior - Single Topology/Address Family/Algorithm

The following mapping entries exist in the database. For brevity, Topology/Algorithm is omitted and assumed to be (0,0) in all entries.

1. (192, 192.0.2.1/32, 100, 1)
2. (192, 192.0.2.101/32, 200, 1)
3. (128, 192.0.2.1/32, 400, 255) !Prefix conflict with entries 1 and 2
4. (128, 198.51.100.40/32, 200,1) !SID conflict with entry 2

The table below shows what mapping entries will be used in the forwarding plane (Active) and which ones will not be used (Inactive)

Active Entries	Inactive Entries
(192,192.0.2.1/32,100,1)	(128,198.51.100.40/32,200,1)


```

| (192,192.0.2.101/32,200,1) |*(128,192.0.2.1/32,400,1) |
|*(128,192.0.2.2/32,401,99) |*(128,192.0.2.101/32,500,1) |
|*(128,192.0.2.102/32,501,154) | |
+-----+

```

* Derived from (128,192.0.2.1/32,400,255)

3.6. Example Behavior - Multiple Topologies

When using a preference rule the order in which conflict resolution is applied has an impact on what entries are Active when entries for multiple topologies (or algorithms) are present. The following mapping entries exist in the database:

1. (192, 192.0.2.1/32, 100, 1, 0, 0) !Topology 0
2. (192, 192.0.2.1/32, 200, 1, 0, 0) !Topology 0, Prefix Conflict with entry #1
3. (192, 198.51.100.40/32, 200,1,1,0) ! Topology 1, SID conflict with entry 2

The table below shows what mapping entries will be used in the forwarding plane (Active) and which ones will not be used (Inactive) based on the order in which conflict resolution is applied.

Order	Active Entries	Inactive Entries
Prefix-Conflict-First	(192,192.0.2.1/32,100,1,0,0) (192,198.51.100.40/32,200,1,1,0)	(192,192.0.2.101/32,200,1,0)
SID-Conflict-First	(192,192.0.2.1/32,100,1,0,0)	(192,192.0.2.101/32,200,1,0) (192,198.51.100.40/32,200,1,1,0)

This illustrates the advantage of evaluating prefix conflicts within a given topology (or algorithm) before evaluating topology (or algorithm) independent SID conflicts. It insures that entries which will be excluded based on intratopology preference will not prevent a SID assigned in another topology from being considered Active.

[3.7.](#) Guaranteeing Database Consistency

In order to obtain consistent active entries all nodes in a network MUST have the same mapping entry database. Mapping entries can be obtained from a variety of sources.

- o SIDs can be configured locally for prefixes assigned to interfaces on the router itself. Only SIDs which are advertised to protocol peers can be considered as part of the mapping entry database.
- o SIDs can be received in prefix reachability advertisements from protocol peers. These advertisements may originate from peers local to the area or be leaked from other areas and/or redistributed from other routing protocols.
- o SIDs can be received from SRMS advertisements - these advertisements can originate from routers local to the area or leaked from other areas
- o In cases where multiple routing protocols are in use mapping entries advertised by all routing protocols MUST be included.

[3.8.](#) Minimizing the occurrence of conflicts

Conflicts in SID advertisements are always the result of a misconfiguration. Conflicts may occur either in the set of advertisements originated by a single node or between advertisements originated by different nodes.

Conflicts which occur within the set of advertisements (PFX and SRMS) originated by a single node SHOULD be prevented by configuration validation on the originating node.

It is possible to minimize the occurrence of conflicts between advertisements originated by different routers if new configuration is validated against the current state of the conflict resolution database before the configuration is advertised. How this is done is an implementation issue which is out of scope of this document.

[4.](#) Scope of SR-MPLS SID Conflicts

The previous section defines the types of SID conflicts and procedures to resolve such conflicts when using an MPLS dataplane. The mapping entry database used MUST be populated with entries for destinations for which the associated SID will be used to derive the labels installed in the forwarding plane of routers in the network. This consists of entries associated with intra-domain routes.

There are cases where destinations which are external to the domain are advertised by protocol speakers running within that network - and it is possible that those advertisements have SIDs associated with those destinations. However, if reachability to a destination is topologically outside the forwarding domain of the protocol instance then the SIDs for such destinations will never be installed in the forwarding plane of any router within the domain - so such advertisements cannot create a SID conflict within the domain. Such entries therefore MUST NOT be installed in the database used for intra-domain conflict resolution.

Consider the case of two sites "A and B" associated with a given [\[RFC4364\]](#) VPN. Connectivity between the sites is via a provider backbone. SIDs associated with destinations in Site A will never be installed in the forwarding plane of routers in Site B. Reachability between the sites (assuming SR is being used across the backbone) only requires using a SID associated with a gateway PE. So a destination in Site A MAY use the same SID as a destination in Site B without introducing any conflict in the forwarding plane of routers in Site A.

Such cases are handled by insuring that the mapping entries in the database used by the procedures defined in the previous section only include entries associated with advertisements within the site.

5. Conflict Resolution and non-forwarding nodes

The previous sections define conflict resolution behavior required of nodes which perform forwarding. But conflict resolution also impacts other entities e.g., controllers. If a controller were to define an explicit path using a SID in a way that is inconsistent with the set of Active entries produced by conflict resolution procedures used by the forwarding nodes then traffic following the explicit path may be misdelivered.

To prevent this such an entity MUST either implement the conflict resolution procedures defined above or implement an alternate form of conflict resolution which produces a subset of the Active entries which result from the conflict resolution procedures defined above. One such alternate form is to consider Inactive any mapping entry

which has either a prefix conflict or a SID conflict with any other mapping entry.

6. Security Considerations

The ability to introduce SID conflicts into a deployment may compromise traffic forwarding. Protocol specific security mechanisms

Ginsberg, et al. Expires January 3, 2018 [Page 17]

Internet-Draft sr-conflict-resolution July 2017

SHOULD be used to insure that all SID advertisements originate from trusted sources.

7. IANA Consideration

This document has no actions for IANA.

8. Acknowledgements

The authors would like to thank Jeff Tantsura, Wim Henderickx, Bruno Decraene, and Stephane Litkowski for their careful review and content suggestions.

9. References

9.1. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", [BCP 14](#), [RFC 2119](#), DOI 10.17487/RFC2119, March 1997, <<http://www.rfc-editor.org/info/rfc2119>>.
- [RFC4364] Rosen, E. and Y. Rekhter, "BGP/MPLS IP Virtual Private Networks (VPNs)", [RFC 4364](#), DOI 10.17487/RFC4364, February 2006, <<http://www.rfc-editor.org/info/rfc4364>>.
- [SR-BGP] "Segment Routing Prefix SID extensions for BGP, [draft-ietf-idr-bgp-prefix-sid-06](#)(work in progress)", June 2017.
- [SR-IS-IS] "IS-IS Extensions for Segment Routing, [draft-ietf-isis-segment-routing-extensions-13](#)(work in progress)", June 2017.

- [SR-MPLS] "Segment Routing with MPLS dataplane, [draft-ietf-spring-segment-routing-mpls-10](#)(work in progress)", June 2017.
- [SR-OSPF] "OSPF Extensions for Segment Routing, [draft-ietf-ospf-segment-routing-extensions-17](#)(work in progress)", June 2017.
- [SR-OSPFv3] "OSPFv3 Extensions for Segment Routing, [draft-ietf-ospf-ospfv3-segment-routing-extensions-09](#)(work in progress)", March 2017.

Ginsberg, et al. Expires January 3, 2018 [Page 18]

Internet-Draft sr-conflict-resolution July 2017

[9.2.](#) Informational References

- [SR-ARCH] "Segment Routing Architecture, [draft-ietf-spring-segment-routing-12](#)(work in progress)", June 2017.

[Appendix A.](#) Alternative SID Conflict Resolution Policy Discussion

A number of approaches to resolving SID conflicts were considered during the writing of this document. Two general approaches with a total of three policy alternatives were considered. This Appendix documents the alternatives considered. All content in this section is non-normative.

Two general approaches can be used to process conflicting entries.

1. Conflicting entries can be ignored
2. A standard preference algorithm can be used to choose which of the conflicting entries will be used

The following sections discuss these two approaches in more detail.

[A.1.](#) Policy: Ignore conflicting entries

In cases where entries are in conflict none of the conflicting entries are used i.e., the network operates as if the conflicting

advertisements were not present.

Implementations are required to identify the conflicting entries and ensure that they are not used.

[A.2.](#) Policy: Preference Algorithm/Quarantine

For entries which are in conflict properties of the conflicting advertisements are used to determine which of the conflicting entries are used in forwarding and which are "quarantined" and not used. Losing mapping entries with ranges greater than 1 are quarantined in their entirety.

This approach requires that conflicting entries first be identified and then evaluated based on a preference rule. Based on which entry is preferred this in turn may impact what other entries are considered in conflict i.e. if A conflicts with B and B conflicts with C - it is possible that A does NOT conflict with C. Hence if as a result of the evaluation of the conflict between A and B, entry B is not used the conflict between B and C will not be detected.

Ginsberg, et al.

Expires January 3, 2018

[Page 19]

Internet-Draft

sr-conflict-resolution

July 2017

[A.3.](#) Policy: Preference algorithm/ignore overlap only

A variation of the preference algorithm approach when applied to mapping entries with ranges greater than 1 is to quarantine only the portions of the less preferred entry which actually conflict. The original entry is logically considered as a set of entries with a range of 1, each of which inherits the range value of the original entry for purposes of applying the preference rule.

[A.4.](#) Example Behavior - Single Topology/Address Family/Algorithm

The following mapping entries exist in the database. For brevity, Topology/Algorithm is omitted and assumed to be (0,0) in all entries.

1. (192, 192.0.2.1/32, 100, 1)
2. (192, 192.0.2.101/32, 200, 1)
3. (128, 192.0.2.1/32, 400, 255) !Prefix conflict with entries 1 and

4. (128, 198.51.100.40/32, 200,1) !SID conflict with entry 2

The table below shows what mapping entries will be used in the forwarding plane (Active) and which ones will not be used (Inactive) under the three candidate policies:

Policy	Active Entries	Inactive Entries
Ignore		(192,192.0.2.1/32,100,1) (192,192.0.2.101/32,200,1) (128,192.0.2.1/32,400,255) (128,198.51.100.40/32,200,1)
Quarantine	(192,192.0.1.1/32,100,1) (192,192.0.2.101/32,200,1)	(128,192.0.2.1/32,400,255) (128,198.51.100.40/32,200,1)
Ignore-Overlap-Only	(192,192.0.2.1/32,100,1) (192,192.0.2.101/32,200,1) *(128,192.0.2.2/32,401,99) *(128,192.0.2.102/32,501,153)	(128,198.51.100.40/32,200,1) *(128,192.0.2.1/32,400,1) *(128,192.0.2.101/32,500,1)

* Derived from (128,192.0.2.1/32,400,300)

[A.5.](#) Evaluation of Policy Alternatives

The previous sections have defined three alternatives for resolving conflicts - ignore, quarantine, and ignore overlap-only.

The ignore policy impacts the greatest number of mapping entries as all prefix/SID pairs contained in an advertisement which has a conflict are considered Inactive.

Quarantine allows forwarding for some destinations which have a conflict to be supported - but losing mapping entries with ranges greater than 1 are declared Inactive in their entirety. This may

result in not using individual prefix/SID entries contained within the quarantined advertisement which do not have a conflict.

Ignore-overlap-only maximizes the entries which may be Active as each prefix/SID pair contained within an advertised mapping entry with range greater than 1 is evaluated independent of the other entries within the same advertisement. To implement this alternative advertised mapping entries with a range greater than 1 which have a conflict with other advertised mapping entries have to logically be split into 2 or more "derived mapping entries". The derived mapping entries then fall into two categories - those that are in conflict with other mapping entries and have lost based on the preference rule and those which are either NOT in conflict or have won based on the preference rule. The former are considered Inactive while the latter are considered Active. Each time the underived mapping database is updated the derived entries have to be recomputed based on the updated database. Internal data structures have to be maintained which maintain the relationship between the advertised mapping entry and the set of derived mapping entries. All nodes in the network have to achieve the same behavior regardless of implementation internals.

There is then a tradeoff between a goal of maximizing advertised mapping entry usage and the risks associated with increased implementation complexity.

Consensus of the working group is that maximizing the use of the advertised prefix/SID pairs is the most important deployment consideration - therefore ignore-overlap-only has been specified as the standard policy which MUST be implemented by all nodes which support SR-MPLS.

Ginsberg, et al.

Expires January 3, 2018

[Page 21]

Internet-Draft

sr-conflict-resolution

July 2017

Authors' Addresses

Les Ginsberg
Cisco Systems
821 Alder Drive

Milpitas, CA 95035
USA

Email: ginsberg@cisco.com

Peter Psenak
Cisco Systems
Apollo Business Center Mlynske nivy 43
Bratislava 821 09
Slovakia

Email: ppsenak@cisco.com

Stefano Previdi
Cisco Systems

Email: stefano@previdi.net

Martin Pilka

Email: martin@infobox.sk