SPRING Working Group
Internet-Draft
Intended status: Standards Track
Expires: July 8, 2021

Shaofu. Peng
ZTE Corporation

January 4, 2021

Truncated SID Inter Domain Considerations
draft-peng-spring-truncated-sid-inter-domain-00

Abstract

This document introduces a method for interworking between domains of Segment Routing in IPv6 network that use different levels of Segment Identifier's compression.

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 https://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 July 8, 2021.

Copyright Notice

Copyright (c) 2021 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 (https://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 2
2. Conventions used in this document 3
2.1. Requirements Language 3
3. Interworking Between Different Domains 3
4. Control Plane Extensions 5
4.1. Extensions to BGP-LS 5
4.2. Extensions to IGP 5
4.2.1. Extensions to IS-IS 5
4.2.2. Extensions to OSPFv3 6
5. Security Considerations 7
6. IANA Considerations 7
6.1. BGP-LS SRv6 SID Swapped Block Sub-TLV 7
6.2. IS-IS SRv6 SID Swapped Block Sub-TLV 7
6.3. OSPFv3 SRv6 SID Swapped Block Sub-TLV 7
7. References 7
7.1. Normative References 7
7.2. Informative References 8
Author's Address 9

1. Introduction

Segment Routing [RFC8402] leverages the source routing paradigm. An ingress node steers a packet through an ordered list of instructions, called segments. [I-D.ietf-spring-srv6-network-programming] proposes the Segment Routing in IPv6 (SRv6) Network Programming to specify a packet processing program by encoding a sequence of instructions in the IPv6 packet's extension header.
[I-D.ietf-spring-srv6-network-programming] defined a Segment Identifier (SID) as consisting of LOC:FUNCT:ARG, where a locator (LOC) is encoded in the L most significant bits of the SID, followed by F bits of function (FUNCT) and A bits of arguments (ARG). A locator may be represented as B:N where B is the SRv6 SID block (IPv6 prefix allocated for SRv6 SIDs by the operator) and N is the identifier of the parent node instantiating the SID. Note that "SID Locator Block" and "BLOCK" are used interchangeably throughout the document.

As demonstrated in several proposals, such as
[I-D.cl-spring-generalized-srv6-for-cmpr] and
[I-D.mirsky-6man-unified-id-sr], SIDs could be allocated by SRv6 nodes from within the same SRv6 SID Locator Block, and referring to the idea described in [RFC6554] SRH only needs to store the difference between SIDs, such as NODE:FUNCT:ARGS, the non-BLOCK part, and does not need to contain the SRv6 SID Locator Block information. In this case, some or all SIDs encoded in the SRH are the non-BLOCK part of the 128-bit SIDs.

An SRv6 path could traverse domains that use different block prefixes and achieving different compressed SIDs lengths. This document provides a method of higher compression efficiency for interworking between domains. It introduces a new SID flavor and defines the required control plane extensions.

2. Conventions used in this document

2.1. Requirements Language

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. Interworking Between Different Domains

The interworking of different domains is illustrated in Figure 1.

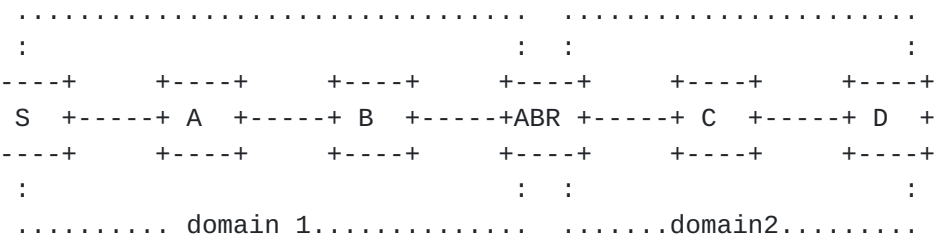


Figure 1: Interworking between domains

An end-to-end SRv6 path from A to D passes through two domains. The block prefix of Domain 1 is BLOCK1, and the compressed SID's length is 32 bits. The block prefix of Domain 2 is BLOCK2, and the compressed SID's length is 16 bits.

The following notation is used in the document, 128bits-BLOCK1-SID-A-32 means that this is a 128-bit SID of node A whose block is BLOCK1 and itself can be compressed to 32 bits if needed. The original segment list before the compression is <128bits-BLOCK1-SID-A-32, 128bits-BLOCK1-SID-B-32, 128bits-BLOCK1-SID-ABR-32, 128bits-BLOCK2-SID-C-16, 128bits-BLOCK2-SID-D-16>.

This segment list is selected/computed by a Path Computation Element (PCE) system or a controller and the details are for further discussion.

A new SID flavor is defined in this document to make the entire segment list more compressible.

Block Swapping (BS) Flavor: it indicates that SRv6 SID Locator Block will be swapped. And the information about the new block with block's length is maintained in the local SID entry.

So in ABR, a 128bits-BLOCK1-ABR-32 SID with BS flavor is allocated. It is within 128-bit SID with block1 and has BS Flavor, and the local SID entry will swap BLOCK1 to BLOCK2. Thus, the next short 16 bit SID can be stitched with BLOCK2 to generate a full 128 bit SID. With the help of the BS flavor, the corresponding compressed segment list could be <32bits-SID-A,32bits-SID-B,32bits-SID-ABR(BS), 16bits-SID-C, 16bits-SID-D>.

32bits-SID-A means this is a 32-bit compressed SID of node A.

32bits-SID-ABR(BS) means this is a 32-bit compressed SID of node ABR with BS flavor.

At the headend S, the packet sent to node A is (SA,DA=BLOCK1+32bits-SID-A)(16bits-SID-D, 16bits-SID-C, 32bits-SID-ABR, 32bits-SID-B, 32bits-SID-A; SL=3), i.e., only one single 128 segment entry is needed in SRH. Note that SL=3 means there are three 32-bits segment items within SRH remaining to be visited.

The packet sent from node A to node B is (SA,DA=BLOCK1+32bits-SID-B)(16bits-SID-D, 16bits-SID-C, 32bits-SID-ABR, 32bits-SID-B, 32bits-SID-A; SL=2).

The packet sent from B to ABR is (SA,DA=BLOCK1+32bits-SID-ABR)(16bits-SID-D, 16bits-SID-C, 32bits-SID-ABR, 32bits-SID-B, 32bits-SID-A; SL=1).

When the packet arrives at ABR, it matches the local SID entry. The endpoint behavior of that entry is END with BS Flavor, and the new block with its block length is provided in the local SID entry. The next short 16 bit SID can be retrieved from SRH and be stitched with the new block to generate a full 128 bit SID to be copied to the DA field. Thus, the packet sent from ABR to C could be (SA,DA=BLOCK2+16bits-SID-C)(16bits-SID-D, 16bits-SID-C, 32bits-SID-ABR, 32bits-SID-B, 32bits-SID-A; SL=1). Note that SL=1 means there are one 16-bits segment items within SRH remaining to be visited.

Generally, multiple BS-flavored SIDs may be allocated on the boundary node. For example, SID-12 with BS flavor indicates that SRv6 SID Locator Block switches from BLOCK1 to BLOCK2, and SID-13 indicates that SRv6 SID Locator Block switches from BLOCK1 to block3.

For the inter-AS scenario, END.X SID allocated for SR-EPE or a direct link can also have BS flavor to indicate the new block information of the next AS.

4. Control Plane Extensions

4.1. Extensions to BGP-LS

This document introduces an SRv6 SID Swapped Block Sub-TLV. It is an optional TLV for use in the BGP-LS Attribute for an SRv6 SID NLRI and as a sub-TLV of the SRv6 End.X, IS-IS SRv6 LAN End.X and OSPFv3 SRv6 LAN End.X TLVs [I-D.ietf-idr-bgpls-srv6-ext]. This TLV is validate only when the SID has a BS flavor.

The TLV has the following format:

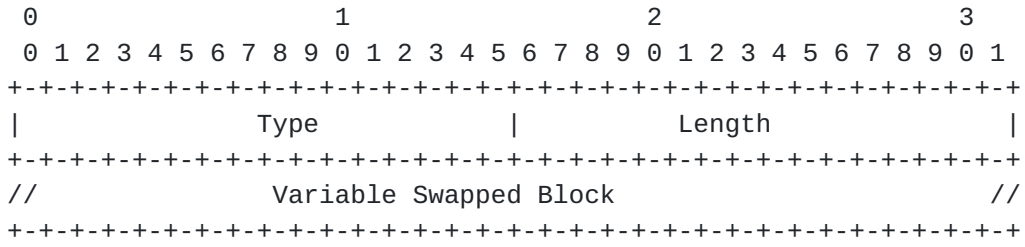


Figure 2: BGP-LS SRv6 SID Swapped Block Sub-TLV

Type: TBA

Length: 2 octet field, indicate the length of Swapped Block.

Swapped Block: This field encodes the advertised swapped block prefix information.

4.2. Extensions to IGP

Different nodes in the same IGP domain may have different block prefixes. So IGP extensions are also needed in this scenario.

4.2.1. Extensions to IS-IS

SID Swapped Block Sub-Sub-TLV is introduced in this document. It is an optional Sub-Sub-TLV of SRv6 End SID Sub-TLV, SRv6 End.X SID Sub-TLV and SRv6 LAN End.X SID Sub-TLV [I-D.ietf-lsr-isis-srv6-extensions].

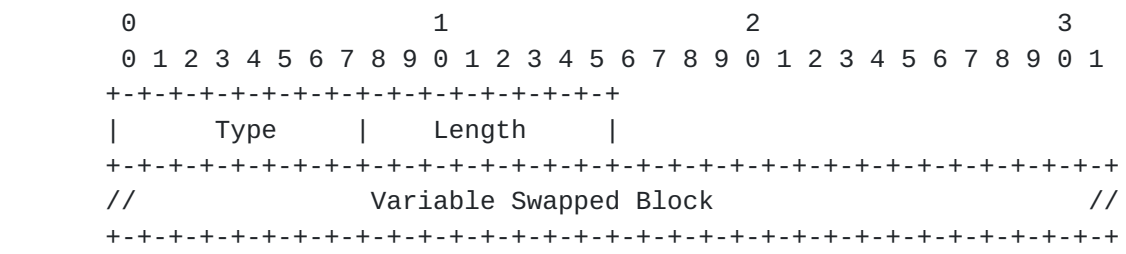


Figure 3: ISIS SRv6 SID Swapped Block Sub-Sub-TLV

Type: TBA, 1 octets.

Length: 1 octet field, indicate the length of Swapped Block.

Swapped Block: This field encodes the advertised swapped block prefix information.

4.2.2. Extensions to OSPFv3

SID Swapped Block Sub-TLV is introduced in this document. It is used as an optional Sub-TLV of SRv6 End SID Sub-TLV, SRv6 End.X SID Sub-TLV, SRv6 LAN End.X SID Sub-TLV [I-D.ietf-lsr-ospfv3-srv6-extensions].

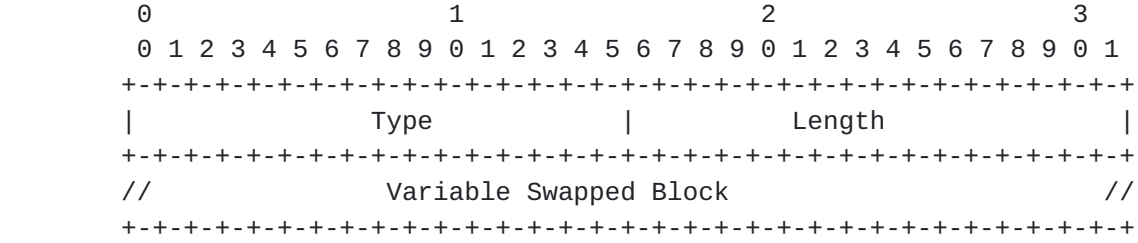


Figure 4: OSPFv3 SRv6 SID Swapped Block Sub-Sub-TLV

Type: TBA, 2 octets.

Length: 2 octet field, indicate the length of Swapped Block.

Swapped Block: This field encodes the advertised swapped block prefix information.

5. Security Considerations

TBD

6. IANA Considerations

6.1. BGP-LS SRv6 SID Swapped Block Sub-TLV

This document requests IANA to assign a new code point in the "BGP-LS Node Descriptor, Link Descriptor, Prefix Descriptor, and Attribute TLVs" sub-registry under the "Border Gateway Protocol - Link State (BGP-LS) Parameters" registry.

Type: TBA

6.2. IS-IS SRv6 SID Swapped Block Sub-TLV

This document requests IANA to assign a new code point in the "Sub-Sub-TLVs for SID Sub-TLVs" registry under the IS-IS registry.

Type: TBA

6.3. OSPFv3 SRv6 SID Swapped Block Sub-TLV

This document requests IANA to assign a new code point in the "OSPFv3 Extended-LSA Sub-TLVs" registry under the "Open Shortest Path First v3 (OSPFv3) Parameters" registry.

Type: TBA

7. References

7.1. Normative References

[I-D.ietf-idr-bgpls-srv6-ext]
Dawra, G., Filsfils, C., Talaulikar, K., Chen, M., daniel.bernier@bell.ca, d., and B. Decraene, "BGP Link State Extensions for SRv6", draft-ietf-idr-bgpls-srv6-ext-05 (work in progress), November 2020.

[I-D.ietf-lsr-isis-srv6-extensions]
Psenak, P., Filsfils, C., Bashandy, A., Decraene, B., and Z. Hu, "IS-IS Extension to Support Segment Routing over IPv6 Dataplane", draft-ietf-lsr-isis-srv6-extensions-11 (work in progress), October 2020.

[I-D.ietf-lsr-ospfv3-srv6-extensions]
Li, Z., Hu, Z., Cheng, D., Talaulikar, K., and P. Psenak, "OSPFv3 Extensions for SRv6", draft-ietf-lsr-ospfv3-srv6-extensions-01 (work in progress), August 2020.

[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>.

[RFC6554] Hui, J., Vasseur, JP., Culler, D., and V. Manral, "An IPv6 Routing Header for Source Routes with the Routing Protocol for Low-Power and Lossy Networks (RPL)", RFC 6554, DOI 10.17487/RFC6554, March 2012, <https://www.rfc-editor.org/info/rfc6554>.

[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>.

[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>.

7.2. Informative References

[I-D.cl-spring-generalized-srv6-for-cmpr]
Cheng, W., Li, Z., Li, C., Clad, F., Aihua, L., Xie, C., Liu, Y., and S. Zadok, "Generalized SRv6 Network Programming for SRv6 Compression", draft-cl-spring-generalized-srv6-for-cmpr-02 (work in progress), November 2020.

[I-D.ietf-spring-srv6-network-programming]
Filsfils, C., Camarillo, P., Leddy, J., Voyer, D., Matsushima, S., and Z. Li, "SRv6 Network Programming", draft-ietf-spring-srv6-network-programming-28 (work in progress), December 2020.

[I-D.mirsky-6man-unified-id-sr]
Cheng, W., Mirsky, G., Peng, S., Aihua, L., and G. Mishra, "Unified Identifier in IPv6 Segment Routing Networks", draft-mirsky-6man-unified-id-sr-07 (work in progress), July 2020.

Author's Address

Peng Shaofu
ZTE Corporation

Email: peng.shaofu@zte.com.cn