IDR
Internet-Draft
Intended status: Standards Track
Expires: 16 October 2025

N. Geng
Huawei
D. Li
Tsinghua University
T. Tong
China Unicom
M. Huang
Zhongguancun Laboratory

14 April 2025

BGP Flow Specification for Source Address Validation
draft-geng-idr-flowspec-sav-05

Abstract

BGP FlowSpec reuses BGP route to distribute infrastructure and propagates traffic flow information with filtering actions. This document proposes some extensions to BGP FlowSpec for disseminating SAV rules.

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 16 October 2025.

Copyright Notice

Table of Contents

1. Introduction

Source Address Validation (SAV) is an efficient method for preventing source address spoofing-based attacks. SAV rules indicate the valid/invalid incoming interfaces of a specific source IP address or source IP prefix. The rules can be deployed on edge routers, border routers, or aggregation routers for checking the validity of intra-domain and inter-domain packets. For invalid packets, filtering actions can be taken such as block, rate-limit, redirect, and sampling [I-D.huang-savnet-sav-table].

There are many mechanisms that can distributedly generate SAV rules on routers ([RFC2827], [RFC3704], [RFC5210], [RFC8704], and [manrs-antispoofing]). To facilitate flexible SAV management and improve validation accuracy, centralized SAV rule dissemination is also needed [I-D.li-savnet-intra-domain-architecture][I-D.wu-savnet-inter-domain-architecture], which can be a complementary to existing distributed SAV mechanisms.

BGP FlowSpec is a convenient and flexible tool for traffic filtering/controlling ([RFC8955], [RFC8956]). It propagates traffic flow information for different traffic control purposes through the BGP protocol extension. Existing BGP FlowSpec has supported source prefix matching and various traffic filtering actions but does not support binding valid/invalid incoming interfaces to source prefixes. With some extensions, BGP FlowSpec can be used for SAV rule dissemination.

This document defines a new flow specification component named Incoming-Interface-Set. SAV rules can be disseminated through BGP FlowSpec by carrying the new flow specification component together with Source Prefix component (or Source Prefix Group component). Traffic filtering actions of existing BGP FlowSpec can also be carried to specify the actions for the packets failing source address validation. This document also defines a new action which indicates where to install the SAV rules in the data-plane.

BGP FlowSpec with the new extensions can be used to disseminate SAV rules to remote routers, which acts as a supplement of existing SAV mechanisms and help improve SAV accuracy.

1.1. Terminology

SAV: Source address validation

SAV Rule: The rule that indicates the valid/invalid incoming interfaces of a specific source IP address or source IP prefix.

Group Identifier: An ID value that identifies a set of interfaces on the target routers (e.g., all the interfaces connected to customer ASes).

1.2. Requirements Language

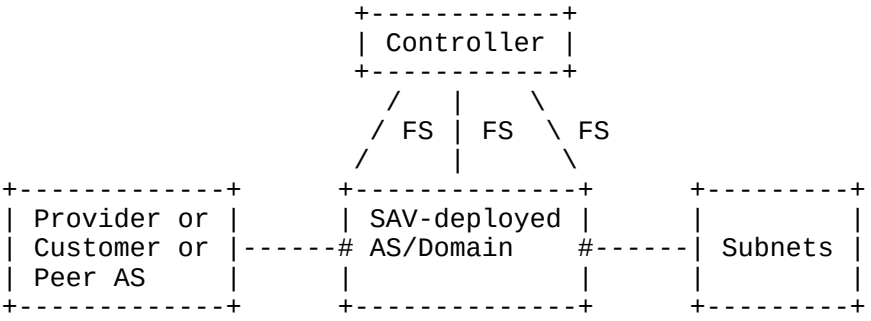
The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

2. BGP FlowSpec for SAV

A SAV rule typically has a format of <source prefix, interface set, validity indicator>. Source prefix is for matching specific packets. Interface set represents a set of physical interfaces from which the packets arrive. Validity indicator indicates whether the packets matching the source prefix and arrival interface are valid or invalid. So, validity indicator has a value of either valid or invalid.

For example, the rule <P1, [intf1, intf2], valid> means the source prefix P1 must arrive the router at interface Intf1 or Intf2, otherwise, P1 is invalid. For the packets with invalid source addresses/prefixes, the filtering actions, such as block, rate-limit, and redirect, can be taken [I-D.huang-savnet-sav-table].

SAV rules can be disseminated to Edge/Border/Aggregation routers (i.e., target routers) through BGP FlowSpec, as shown in the figure below. The controller is used to set up BGP connection with the routers in a SAV-deployed AS or domain.



Existing BGP FlowSpec has supported source prefix matching and various traffic filtering actions. There are also some proposals that group source prefixes by AS numbers ([I-D.wang-idr-flowspec-dip-origin-as-filter]) or community values for good scalability and efficiency. An AS number or a community value can represent a set of source prefixes. For simplicity, the term of Source Prefix Group component will be used to represent such source prefix groups in BGP FlowSpec.

However, existing BGP FlowSpec does not support binding valid/invalid incoming interfaces to source prefixes. Besides, BGP FlowSpec rules are mostly installed in ACL table/firewall table. In contrast, SAV rules may be installed in differect positions of data-plane, such as ACL/firewall, FIB, or independent SAV table. Some extensions are needed by BGP FlowSpec for efficient SAV rule dissemination.

3. Extensions to BGP FlowSpec

3.1. Incoming-Interface-Set Component

3.1.1. Interface Set

To facilitate scalability, the interface set in SAV rules can be grouped. For example, the interfaces can be grouped as:

* Subnet interface set that contains the interfaces connecting a target subnet.

* All customer AS interfaces set or the customer AS interfaces set of a customer AS.

* All lateral peer AS interfaces set or the lateral peer AS interfaces set of a lateral peer AS.

* All transit provider AS interfaces set or the transit provider AS interfaces set of a transit provider AS.

These interface set can be identified by a Group Identifier for easy management. A Group Identifier can have either local meaning or global meaning. On the one hand, it can be a local interface property on the target routers, and the meaning of it depends on the configurations of network administrator. On the other hand, a global meaning Group Identifier field carries AS number, which represents all the interfaces connected to the neighboring AS with the AS number.

Any interface may be associated with one or more Group Identifiers.

3.1.2. Incoming-Interface-Set Encoding

The new flow specification component is encoded in the BGP Flowspec NLRI. It appears together with Source Prefix component or Source Prefix Group component.

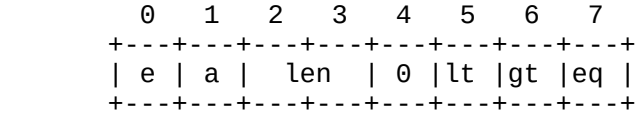
The following new component type is defined:

* Type TBD1: Incoming-Interface-Set

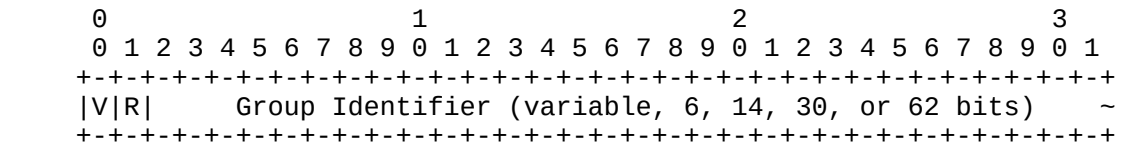
* Encoding: <type (1 octet), [numeric_op, value]+>

The new component contains a set of {numeric_op, value} pairs that are used to match the Incoming-Interface-Set (i.e., the valid or invalid interfaces of a specific source prefix).

The numeric operator (numeric_op) is encoded as (see RFC8955 sec. 4.2.1.1):



The value field is encoded as:



The length of the value field can be 1, 2, 4, or 8 octets, which depends on the len in numeric_op. Particularly, the most two significant bits in the value field are two flags:

* Flag V (1 bit): The most significant bit in the value field. If unset, it means the interface set is invalid for the source prefix. If set, the identified interface set is valid for the source prefix.

* Flag G (1 bit): The second most significant bit in the value field. If unset, the Group Identifier has a local meaning. If set, the Group Identifier has a global meaning, i.e., the Group Identifier field stores an AS number.

Zero Group Identifier (i.e., Group Identifier equaling 0) is a reserved value and does not need to be configured. A NLRI may carry one zero Group Identifier and several non-zero Group Identifiers. The zero Group Identifier means any other interfaces on the target router except the interfaces indicated by non-zero Group Identifiers in the same NLRI. If a NLRI only contains a zero Group Identifier and has no non-zero Group Identifiers, the zero Group Identifier will represent all interfaces on the target router. A NLRI MUST not contain more than one zero Group Identifiers, otherwise, the whole NLRI will be ignored.

The bits lt, gt, and eq can be combined to match a specific Group Identifier or a range of Group Identifiers (e.g., greater than Group ID1 and less than Group ID2). For a range of Group Identifiers, their corresponding flags (i.e., V and R) MUST be the same. Otherwise, the whole NLRI will be ignored.

If a receiving BGP speaker cannot support this new flow specification component type, it MUST discard the NLRI value field that contains such unknown components (section 10 of [RFC8955]). A NLRI value field MUST only contain a Source Prefix component (or Source Prefix Group component) and an Incoming-Interface-Set component. If the NLRI value does not satisfy this principle, the receiving BGP speaker SHOULD discard the NLRI value field (see Section Section 4). Since the NLRI field encoding (Section 4 of [RFC8955]) is defined in the form of a 2-tuple <length, NLRI value>, message decoding can skip over the unknown NLRI value and continue with subsequent remaining NLRIs.

3.1.3. Example

Example: A Flow Specification NLRI encoding for "incoming interfaces {Group ID range [1, 20]} are valid for the packets from 203.0.113.0/24, and other local interfaces are invalid for the source prefix".

Length	Source	Flags+Group Identifier
0c	02 18 cb 00 71	TBD1 03 81 45 94 81 00

Decoded:

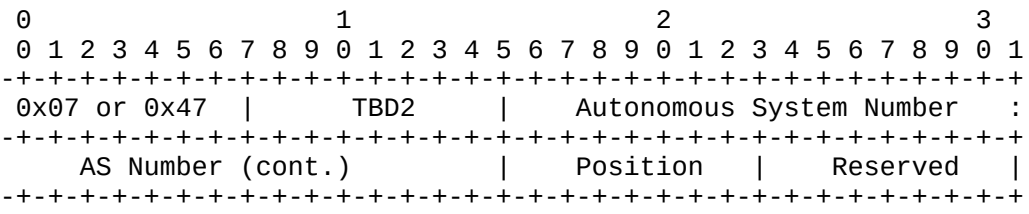
Value		
0x0c	length	12 octets (if len<240, 1 octet)
0x02	type	Type 2 - Source Prefix
0x18	length	24 bit
0xcb	prefix	203
0x00	prefix	0
0x71	prefix	113
TBD1	type	Type TBD1 - Incoming-interface-set
0x03	numeric_op	value size=1, >=
0x81	value	V=1, R=0, ID=1
0x45	numeric_op	"AND", value size=1, <=
0x94	value	V=1, R=0, ID=20
0x81	numeric_op	end-of-list, value size=1, ==
0x00	value	V=0, R=0, ID=0

This constitutes an NLRI with an NLRI length of 12 octets.

3.2. Rule-Position Extended Community

This document proposes a new BGP Route Target extended community called the "rule-position". This document expands the definition of the Route Target extended community to allow a new value of high order octet (Type field) to be 0x07 for the transitive flowspec interface-set extended community, or 0x47 for the non-transitive flowspec interface-set extended community. These are in addition to the values specified in [RFC4360].

The Rule-Position extended community is encoded as follows:



The Position field indicates that where to install the SAV rule. Some values can be:

* 0: ACL

* 1: FIB

* 2: Independent SAV table

* Other values: reserved

Since FIB and Independent SAV table can only match address prefix and interface, the Rule-Position extended community with the Position field equaling 1 or 2 means the NLRI is specific for SAV rule dissemination, and the unrelated components (not Source Prefix (Group) or Incoming-Interface-Set) SHOULD NOT appear in the NLRI.

4. General Usages

The Incoming-Interface-Set component can be used as a general flow specification instead of SAV-specific component. Other components can be combined with the new component for matching specific traffic.

5. Error Handling

TBD

6. IANA Considerations

6.1. Incoming-Interface-Set Component

This document requests a new entry in "Flow Spec component types registry" with the following values:

Type	IPv4/IPv6 Name	Reference
TBD1	Incoming-Interface-set	This document

6.2. Rule-Position Extended Community

This document requests a new subtype TBD2 within Transitive and Non-Transitive Extended Communities. This sub-type shall be named "rule-position", with a reference to this document.

7. Security Considerations

TBD.

8. Acknowledgements

Many thanks to the comments from Shunwan Zhuang, Susan Hares, Jeffrey Haas, Mingxing Liu, etc.

9. References

9.1. Normative References

[I-D.wang-idr-flowspec-dip-origin-as-filter]
Wang, H., Wang, A., and S. Zhuang, "Destination-IP-Origin-AS Filter for BGP Flow Specification", Work in Progress, Internet-Draft, draft-wang-idr-flowspec-dip-origin-as-filter-10, 13 January 2025, <https://datatracker.ietf.org/doc/html/draft-wang-idr-flowspec-dip-origin-as-filter-10>.

[RFC8955] Loibl, C., Hares, S., Raszuk, R., McPherson, D., and M. Bacher, "Dissemination of Flow Specification Rules", RFC 8955, DOI 10.17487/RFC8955, December 2020, <https://www.rfc-editor.org/info/rfc8955>.

[RFC8956] Loibl, C., Ed., Raszuk, R., Ed., and S. Hares, Ed., "Dissemination of Flow Specification Rules for IPv6", RFC 8956, DOI 10.17487/RFC8956, December 2020, <https://www.rfc-editor.org/info/rfc8956>.

[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>.

[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>.

9.2. Informative References

[I-D.li-savnet-intra-domain-architecture]
Li, D., Wu, J., Qin, L., Geng, N., Chen, L., Huang, M., and F. Gao, "Intra-domain Source Address Validation (SAVNET) Architecture", Work in Progress, Internet-Draft, draft-li-savnet-intra-domain-architecture-07, 16 March 2024, <https://datatracker.ietf.org/doc/html/draft-li-savnet-intra-domain-architecture-07>.

[I-D.wu-savnet-inter-domain-architecture]
Li, D., Chen, L., Geng, N., Liu, L., and L. Qin, "Inter-domain Source Address Validation (SAVNET) Architecture", Work in Progress, Internet-Draft, draft-wu-savnet-inter-domain-architecture-12, 3 November 2024, <https://datatracker.ietf.org/doc/html/draft-wu-savnet-inter-domain-architecture-12>.

[I-D.huang-savnet-sav-table]
Huang, M., Cheng, W., Li, D., Geng, N., Liu, Chen, L., and C. Lin, "General Source Address Validation Capabilities", Work in Progress, Internet-Draft, draft-huang-savnet-sav-table-08, 10 December 2024, <https://datatracker.ietf.org/doc/html/draft-huang-savnet-sav-table-08>.

[manrs-antispoofing]
"MANRS Implementation Guide", January 2023, <https://www.manrs.org/netops/guide/antispoofing>.

[RFC2827] Ferguson, P. and D. Senie, "Network Ingress Filtering: Defeating Denial of Service Attacks which employ IP Source Address Spoofing", BCP 38, RFC 2827, DOI 10.17487/RFC2827, May 2000, <https://www.rfc-editor.org/info/rfc2827>.

[RFC3704] Baker, F. and P. Savola, "Ingress Filtering for Multihomed Networks", BCP 84, RFC 3704, DOI 10.17487/RFC3704, March 2004, <https://www.rfc-editor.org/info/rfc3704>.

[RFC5210] Wu, J., Bi, J., Li, X., Ren, G., Xu, K., and M. Williams, "A Source Address Validation Architecture (SAVA) Testbed and Deployment Experience", RFC 5210, DOI 10.17487/RFC5210, June 2008, <https://www.rfc-editor.org/info/rfc5210>.

[RFC8704] Sriram, K., Montgomery, D., and J. Haas, "Enhanced Feasible-Path Unicast Reverse Path Forwarding", BCP 84, RFC 8704, DOI 10.17487/RFC8704, February 2020, <https://www.rfc-editor.org/info/rfc8704>.

Authors' Addresses

Nan Geng
Huawei
Beijing
China
Email: gengnan@huawei.com

Dan Li
Tsinghua University
Beijing
China
Email: tolidan@tsinghua.edu.cn

Tian Tong
China Unicom
Beijing
China
Email: tongt5@chinaunicom.cn

Mingqing Huang
Zhongguancun Laboratory
Beijing
China
Email: huangmq@mail.zgclab.edu.cn