Inter-Domain Routing
Internet-Draft
Updates: 8955 (if approved)
Intended status: Standards Track
Expires: 11 October 2021

J. Haas
Juniper Networks

9 April 2021

BGP Flowspec Capability Bits
draft-haas-flowspec-capability-bits-02

Abstract

BGP Flowspec (RFC 8955) provides the ability to filter traffic using various matching components. The NLRI format currently defined does not permit incremental deployment of new BGP Flowspec components. This draft defines a new BGP Capability to permit incremental deployment of such new Flowspec component types.

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 BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

BGP Flowspec [RFC8955] provides a mechanism to distribute traffic flow spcifications into BGP. One general purpose of these flow specifications is for the distribution of firewall rules to receiving routers, particularly for mitigating distributed denial of service (DDoS) attacks. The flow specification rules are encoded as BGP NLRI [RFC4271].

The matching components of a flow specification NLRI is a serialized set of optional components. The components are documented in [RFC8955], Section 3. [RFC8956] defines IPv6-specific components. The full set of Flowspec component types is maintained in an IANA registry located at the IANA Flow Spec Component Types registry (https://www.iana.org/assignments/flow-spec/flow-spec.xhtml).

Unknown Flowspec component types require treatment as a malformed NLRI ([RFC8955], Section 4.2). This is due to the lack of a mandatory length element for the components in the NLRI. Without such a length, it is not possible to determine how to properly decode unknown components in the Flowspec NLRI.

There has been active interest in the IDR Working Group to extend BGP Flowspec for additional purposes. However, with this difficulty in being able to handle unknown components, those new features are unable to be deployed in a BGP Flowspec domain in an incremental fashion. Either a carefully managed "flag day" deployment is required to avoid disrupting existing sessions, or the Flowspec domain is carefully managed such that devices with incompatible sets of known/unknown components are carefully separated in a "ships in the night" scenario. Both options are fragile and operationally cumbersome.

Some initial discussion has begun for a version 2 of Flowspec in [I-D.hares-idr-flowspec-v2]. That document may eventually address this incremental deployment issue, along with a number of other items.

This document proposes to address the issues of incremental deployment of new BGP Flowspec component types via a new BGP Capability [RFC5492], the BGP Flowpec Capability Bits.

2. BGP Flowspec Capability Bits

BGP Flowspec component types are one octet in length with values in the range from 0..255. The BGP Flowspec Capability Bits encode a bit-string where each supported component type has its respective bit set when the BGP Speaker is willing to receive BGP Flowspec NLRI that contain that component type.

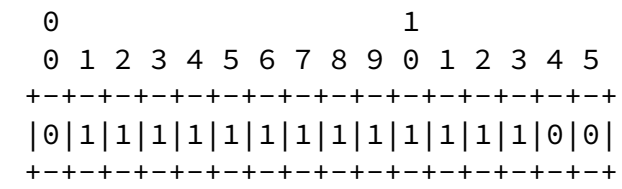
The BGP Flowspec Capability Bits Capability is encoded as follows:

* Capability Code of (TBD).

* Capability Length of 1..32.

* Capability Value contains a bit-string where a bit is set if the underlying BGP Flowspec component is willing to be accepted by BGP Speaker advertising this capability.

Example encoding for Capability Value:



Bit 0 set to 0, bits 1..13 set to 1 showing support for all capabilities for IPv6 Flowspec, bits 14 and 15 are set to 0.

3. Operation

BGP Flowspec Capability Bits not advertised in the encoded bit-string are treated as if they were sent with a value of zero for that bit.

The Capability Length reflects the number of octets it takes to encode the BGP Flowspec Capability Bits. While the total number of octets required to represent the entire range of component types is only 32 octets, implementations SHOULD limit the number of octets transmitted to those required to encode the final one-bit. Space in BGP Capabilities may be limited in some implementations depending on the number of capabilities to be sent. (See [I-D.ietf-idr-ext-opt-param] for discussion on a feature to address this point.)

Bit-values 0 and 255 SHOULD be set to zero as they are RESERVED.

The BGP Flowspec Capability Bits Capability SHOULD be sent by a BGP

Speaker utilizing any AFI/SAFI using BGP Flowspec encoding as defined in [RFC8955], or [RFC8956].

The BGP Flowspec Capability Bits Capability MUST be sent by a BGP Speaker utilizing BGP Flowspec encoding with a component type not defined in those documents previously mentioned. (I.e. component types not in the range 1..13.)

A BGP Speaker that has received the BGP Flowspec Capability Bits Capability MUST NOT originate or propagate a BGP Flowspec encoded NLRI that contains a component types that is not present in the received bit-string.

A BGP Speaker that has received a BGP Flowspec related AFI/SAFI without this Capability MUST treat the absence as equivalent to having received the Capability Bits covered by the base specification for its defining RFC, [RFC8955] or [RFC8956].

4. Propagation of Known Components and Mismatch with Local Filtering Capabilities

There may be circumstances where a BGP Speaker is capable of parsing Flowspec components that it is not capable of implementing as filters. Section 4.2 of [RFC8955] specifies that:

> "All combinations of components within a single Flow Specification are allowed. However, some combinations cannot match any packets (e.g., "ICMP Type AND Port" will never match any packets) and thus SHOULD NOT be propagated by BGP."

This document updates that text to:

> "All combinations of components within a single Flow Specification are allowed. However, some combinations cannot match any packets (e.g., "ICMP Type AND Port" will never match any packets) and thus SHOULD NOT be propagated by BGP.
>
> "When BGP Flowspec component types are understood and the operator determines that deployment-wide filtering intent would not be comporomised by propagating Flowepec routes that cannot match any

packets, it SHOULD propagate the route in BGP. This permits NLRI with known components to be propagated to downstream BGP Speakers in the deployment."

5. BGP Flowspec Implications for Filtered NLRI

BGP Flowspec NLRI encode match operations for traffic filtering rules. Filtering is an ordered operation. Since the current encoding of the NLRI does not supply explicit filtering order, the protocol imposes a forwarding order based on the contents of the NLRI.

When a BGP Flowspec NLRI is not propagated due to filtering by this feature, or by user policy, there is the potential that the network-wide filtering intent may be compromised by the missing rules. The exact impact of this on filtering will depend on the relative independence of the full set of BGP Flowspec routes in the BGP Flowspec routing domain.

Operators must exercise care when deploying BGP Flowspec features with new component types to understand the propagation of such routes in their deployment, and the impact that filtering may have on the routes they wish to originate.

6. Error Handling

If a BGP Speaker implementing this document has transmitted BGP Flowspec Capability Bits to its peer and receives a BGP Flowspec NLRI with an unacceptable component (not in its bit-string), it MAY terminate the BGP session by sending a NOTIFICATION message.

7. Acknowledgements

Thanks to Aseem Choudhary, Jakob Heitz, Christoph Loibl, Robert Raszuk for their comments on this proposal.

8. Security Considerations

All of the Security Considerations for [RFC8955] and [RFC8956] still

apply.

Additionally, the BGP Flowspec Capability Bits may cause implicit filtering of some BGP Flowspec NLRI in a Flowspec domain. Depending on the relative independence of the traffic matched by the BGP Flowspec rules in the ordering required by their specifications, such filtered NLRI may result in impact to the desired domain-wide filtering behaviors.

9. IANA Considerations

IANA is requested to assign a new BGP Capability to the Capability Codes registry from the First Come, First Served pool. The Reference for the registration is this document. The Change Controller is IETF.

10. References

10.1. 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, <https://www.rfc-editor.org/info/rfc2119>.

[RFC5492] Scudder, J. and R. Chandra, "Capabilities Advertisement with BGP-4", RFC 5492, DOI 10.17487/RFC5492, February 2009, <https://www.rfc-editor.org/info/rfc5492>.

[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>.

[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>.

10.2. Informative References

[I-D.hares-idr-flowspec-v2]
Hares, S., "BGP Flow Specification Version 2", Work in Progress, Internet-Draft, draft-hares-idr-flowspec-v2-00, 25 June 2016, <http://www.ietf.org/internet-drafts/draft-hares-idr-flowspec-v2-00.txt>.

[I-D.ietf-idr-ext-opt-param]
Chen, E. and J. Scudder, "Extended Optional Parameters Length for BGP OPEN Message", Work in Progress, Internet-Draft, draft-ietf-idr-ext-opt-param-09, 21 August 2020, <http://www.ietf.org/internet-drafts/draft-ietf-idr-ext-opt-param-09.txt>.

[RFC2578] McCloghrie, K., Ed., Perkins, D., Ed., and J. Schoenwaelder, Ed., "Structure of Management Information Version 2 (SMIv2)", STD 58, RFC 2578, DOI 10.17487/RFC2578, April 1999, <https://www.rfc-editor.org/info/rfc2578>.

[RFC4271] Rekhter, Y., Ed., Li, T., Ed., and S. Hares, Ed., "A Border Gateway Protocol 4 (BGP-4)", RFC 4271, DOI 10.17487/RFC4271, January 2006, <https://www.rfc-editor.org/info/rfc4271>.

Appendix A. Encoding of the Bit-String

IETF has a mixed history in terms of how bit numbering is described. The format as used in this document, where the left-most bit sent on the wire is bit zero, is consistent with IETF PDU diagrams and also the SNMP BITS construct [RFC2578], Section 7.1.4.

That said, the author is aware of how annoying the code for that construct can be.

Appendix B. Open Issues

* Are there circumstances where advertising capability bits for BGP

Flowspec NLRI that need to vary on a per AFI-SAFI basis? Currently, the IANA registry is a single name space for all supported and proposed BGP address families. As an example, the Flowspec for NVO3 feature has components that are defined that do not have incremental deployment issues due to being well formed with a length field. However, since it still includes existing Flowspec filtering for the outer and inner IP headers, the issues addressed by this proposal still apply.

Author's Address

Jeffrey Haas
Juniper Networks

Email: jhaas@juniper.net