Inter-Domain Routing
Internet-Draft
Intended status: Standards Track
Expires: 29 October 2021

J. Haas
Juniper Networks
S. Hares
Hickory Hill Consulting
S. Maduschke
Verizon

27 April 2021

BGP Flowspec Explicit Term Ordering
draft-haas-idr-flowspec-term-order-00

Abstract

BGP Flowspec (RFC 8955) provides a mechanism for matching traffic flows. The ordering of the Flow Specifications defined by that RFC is provided by a sorting function that uses the contents of the received BGP NLRI; that NLRI does not contain an explicit ordering component. The RFC's sorting function permits for origination of Flowspec NLRI from multiple BGP Speakers and is generally appropriate for mitigating distributed denial-of-service (DDoS) attacks.

There are circumstances where the implicit RFC 8955 sorting order is not appropriate. This document defines a mechanism that permits individual Flowspec NLRI to influence their sort order.

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

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 29 October 2021.

Table of Contents

1. Introduction

BGP Flowspec [RFC8955] creates a mechanism for matching traffic flows and taking action upon them. The BGP Flowspec NLRI format defines multiple components that may be used to match such traffic. Traffic may be matched by more than one BGP Flowpec NLRI, either before or after the application of Traffic Filtering Actions (Section 7, [RFC8955]).

[RFC8955] does not provide for a mechanism where the originator of a BGP Flowspec NLRI can influence its processing order. Section 5.1 of [RFC8955] provides for a sorting function on a BGP Speaker defining the processing order of received BGP Flowspec NLRI. That sorting mechanism permits multiple BGP Speakers in a Flowspec domain to originate Flowspec NLRI without coordinating the processing order at a given BGP Speaker.

That sorting order is generally appropriate for mitigating distributed denial-of-service attacks (DDoS). Flow specification rules first match on related destinations, followed by sources, and then later a well-defined set of components. Longer sets of components are considered a better match, and thus "more specific" in many cases.

While this sort order has generally worked well for DDoS mitigation, sometimes the implicit ordering is problematic. Some of these problems are implementation specific: Long rule sets might be better sorted into higher impact filters near the top of the list. Mixtures of rules that are otherwise independent are sorted in such a way that firewall optimizations are not efficiently run.

Some initial discussion has begun for a version 2 of Flowspec in [I-D.hares-idr-flowspec-v2]. Part of that proposal is a mechanism to provide for explicit rule ordering as part of the Flowspec v2 NLRI.

This document proposes an alternative mechanism to provide for such explicit rule ordering with a minor extension to Flowspec v1.

2. Term Order Component Type

2.1. Type 0 - Term Order

Encoding: <type (1 octet), length (1 octet), term order (variable)>

Defines the relative term order for this BGP Flowspec NLRI.

The value of the length MUST be 1, 2, or 4. The length SHOULD be chosen to be the smallest possible value to properly encode the term order value.

2.2. Discussion

The choice of Component Type 0, currently RESERVED by [RFC8955], is intended to be minimally disruptive to the sorting function and deployed code for BGP Flowspec. Consider the following text from Section 5.1 of that RFC:

> "The relative order of two Flow Specifications is determined by comparing their respective components. The algorithm starts by comparing the left-most components (lowest component type value) of the Flow Specifications. If the types differ, the Flow Specification with lowest numeric type value has higher precedence (and thus will match before) than the Flow Specification that doesn't contain that component type. If the component types are the same, then a type-specific comparison is performed (see below). If the types are equal, the algorithm continues with the next component."

By using Component Type 0, the ability to bias sort order is provided without a change to the remaining sorting semantics used by [RFC8955] and other proposals.

3. Operation

The term order value, when present in a BGP Flowspec NLRI, is intended to provide a logical order to that NLRI vs. other NLRI with that component. A lower term order value has a higher precedence than a higher term order value.

A BGP Flowspec NLRI with no term order component is considered to be lower precedence versus a BGP Flowspec NLRI with a term order component. This is consistent with existing BGP Flowspec sorting rules.

The same term order value MAY occur more than once in a set of BGP Flowspec NLRI.

The term order value is not intended to supplant the ordering mechanism for a firewall implementation. Its only purpose is to provide for biasing the sorting of received BGP Flowspec NLRI.

3.1. Incremental Deployment

[I-D.haas-flowspec-capability-bits] is required to deploy this feature for Flowspec v1. When a BGP Speaker wishes to use, originate, or propagate BGP Flowspec NLRI with the term order component, that BGP Speaker MUST advertise the BGP Flowspec Capability Bits with bit 0 set to a value of 1.

4. Error Handling

A BGP Flowpsec Term Order Component with a length that is not 1, 2, or 4 is considered syntactically incorrect per Section 5.3 of [RFC7606]. Upon receiving such syntactically incorrect NLRI, the BGP session SHALL be reset by sending a NOTIFICATION message.

5. Acknowledgements

TBD.

6. Security Considerations

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

This feature provides for the ability to bias the installed filter order of BGP Flow Specification NLRI. The default sort order provided by [RFC8955] serves to cluster rules targeting traffic for a given destination and/or source. By providing an ability to alternatively order such rules, more general rules impacting more traffic may have precedence.

Operators must take sufficient care to ensure that such more general rules are considered systematically in the deployment. This may include the ability to prohibit rules with a term order outside of a specific value range from being accepted.

Operators may wish to prohibit other ASes from originating or propagating BGP Flowspec NLRI with the term order component, even while exercising the Validation Procedures of Section 6 of [RFC8955].

7. IANA Considerations

Upon approval of this document as an RFC, IANA is requested to assign Type Value 0 from the IANA Flow Spec Component Types registry (https://www.iana.org/assignments/flow-spec/flow-spec.xhtml). The IPv4 Name and IPv6 name for Type 0 will be "Term Order". The Reference will be this document.

8. References

8.1. Normative References

[I-D.haas-flowspec-capability-bits]
Haas, J., "BGP Flowspec Capability Bits", Work in Progress, Internet-Draft, draft-haas-flowspec-capability-bits-02, 9 April 2021, <https://www.ietf.org/internet-drafts/draft-haas-flowspec-capability-bits-02.txt>.

[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>.

[RFC7606] Chen, E., Ed., Scudder, J., Ed., Mohapatra, P., and K. Patel, "Revised Error Handling for BGP UPDATE Messages", RFC 7606, DOI 10.17487/RFC7606, August 2015, <https://www.rfc-editor.org/info/rfc7606>.

[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>.

8.2. 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>.

Appendix A. Open Issues

* After sufficient discussion has been given to this proposal, update the python pseudocode example to include interaction with this feature.

Authors' Addresses

Jeffrey Haas
Juniper Networks

Email: jhaas@juniper.net

Susan Hares
Hickory Hill Consulting

Email: shares@ndzh.com

Sven Maduschke
Verizon

Email: sven.maduschke@de.verizon.com