



# BGP Flow Spec for DDoS mitigation



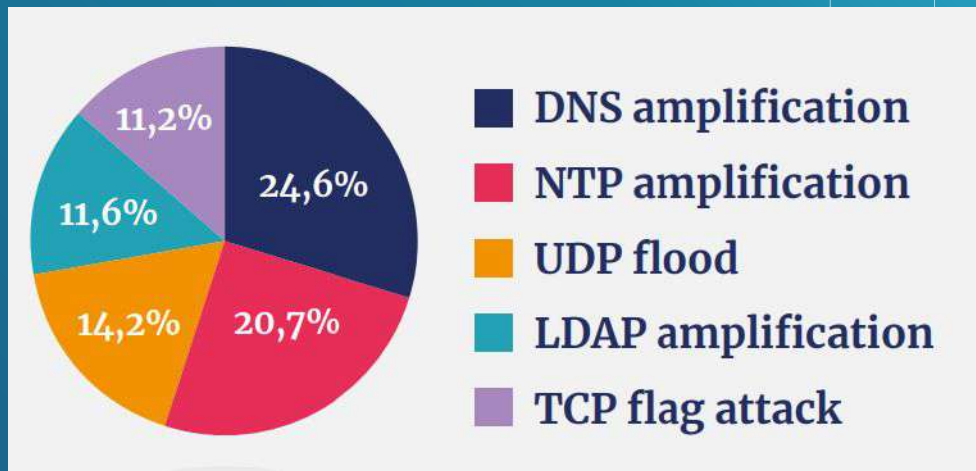
# Hello

I'm Pavel Odintsov, DDoS mitigation enthusiast, the author of FastNetMon.  
<https://fastnetmon.com> and CTO of FastNetMon LTD.

Ways to contact me:

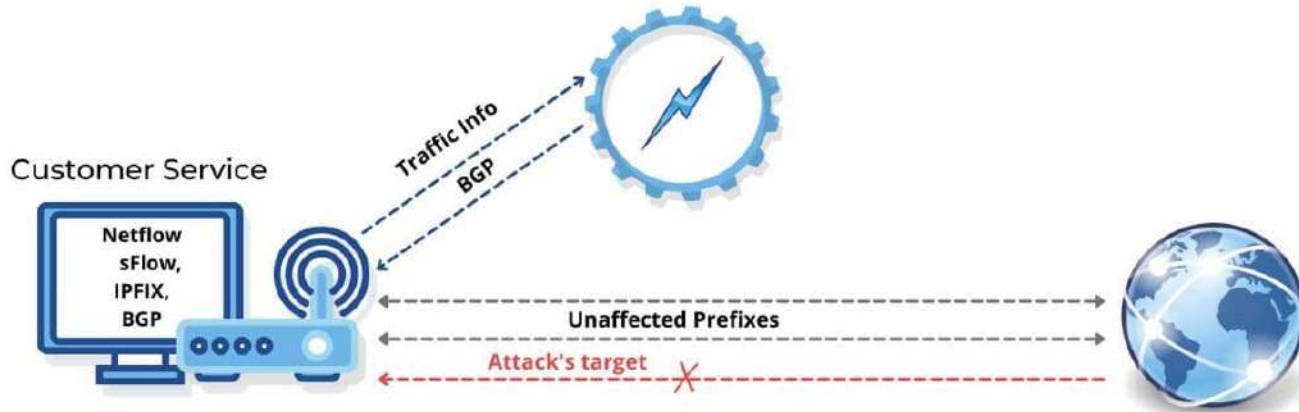
- [linkedin.com/in/podintsov](https://www.linkedin.com/in/podintsov)
- [github.com/pavel-odintsov](https://github.com/pavel-odintsov)
- [twitter.com/odintsov\\_pavel](https://twitter.com/odintsov_pavel)
- IRC, Libera Chat, `pavel_odintsov`
- [pavel@fastnetmon.com](mailto:pavel@fastnetmon.com)

# Current DDoS Weather



Data provided by The Dutch National Scrubbing Center (NaWas), Q2 2022

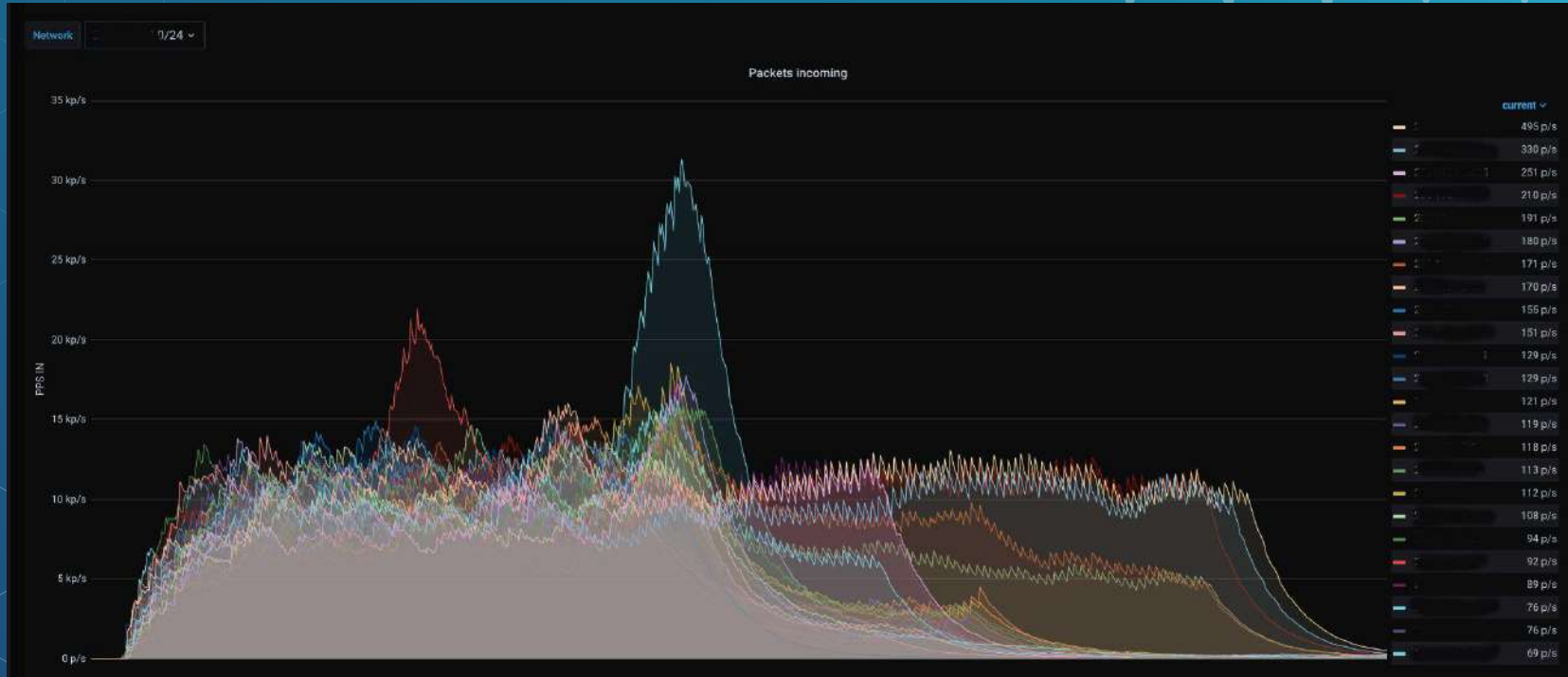
# BGP Blackhole / RTBH



A stylized illustration in light blue outlines on a dark blue background. It features a planet with a ring system and three small circles on its surface, several five-pointed stars of varying sizes, and a rocket ship with a flame trail at the bottom left.

What is the  
problem?

# Carpet Bombing Attack





# What is BGP Flow Spec / RFC5575

- Protocol to configure distributed firewall
- BGP NLRI (Network Layer Reachability Information)
- RFC 5575 standard was published in 2009

# BGP Flow Spec filtering capabilities

- Source prefix (IPv4 or IPv6)
- Destination prefix (IPv4 or IPv6)
- IP Protocol number
- List or range of source ports for TCP and UDP
- List or range of destination ports for TCP and UDP
- ICMP code
- TCP flags
- Packet length
- Fragmentation flags (do not fragment, is fragment, first or last fragment)
- DSCP





# BGP Flow Spec filtering actions

- Drop
- Rate limit
- Accept
- Mark (DSCP)
- Redirect to VRF
- Redirect to nexthop (draft)

# Workgroup spent 6 years on RFC 5575

**Dissemination of Flow Specification Rules**  
draft-ietf-idr-flow-spec-09

Status IESG evaluation record IESG writeups Email expansions History

Versions 09


draft-marques-idr-flow-spec 00  
draft-ietf-idr-flow-spec rk5575


| Version | Date     |
|---------|----------|
| 00      | Jan 2003 |
| 01      | Oct 2004 |
| 02      | Dec 2004 |
| 03      | Aug 2005 |
| 04      | Jul 2007 |
| 05      | Aug 2007 |
| 06      | Apr 2008 |
| 07      | Sep 2008 |
| 08      | Nov 2008 |
| 09      | Jan 2009 |
| 09      | Nov 2009 |
| 09      | Mar 2009 |
| 09      | Aug 2009 |

**The information below is for an old version of the document that is already published as an RFC.**

|                 |  |
|-----------------|--|
| <b>Document</b> | <b>Type</b> RFC Internet-Draft (idr WG)  |
|                 | <b>Authors</b> Pedro Marques sa, Jared Mauch sa, Nischal Sheth sa, Barry Greene sa, Robert Raszuk sa, Danny McPherson sa |
|                 | <b>Last updated</b> 2020-01-21 (latest revision 2009-05-26)  |
|                 | <b>Replaces</b> draft-marques-idr-flow-spec  |
|                 | <b>Stream</b> Internet Engineering Task Force (IETF)   |
|                 | <b>Formats</b> <a href="#">plain text</a> <a href="#">htmlized</a> <a href="#">pdfized</a> <a href="#">bibtext</a>       |
|                 | <b>Reviews</b> SECDIR Last Call Review   |
| <b>Stream</b>   | <b>WG state</b> WG Document  |
|                 | <b>Document shepherd</b> No shepherd assigned  |
| <b>IESG</b>     | <b>IESG state</b> RFC 5575 (Proposed Standard)   |
|                 | <b>Consensus Boilerplate</b> Unknown   |
|                 | <b>Telechat date</b>   |
|                 | <b>Responsible AD</b> Adrian Farrel  |
|                 | <b>Send notices to</b> (None)  |

# Support on Juniper, JunOS 12.3, March 2012?

Border Gateway Protocol (BGP) 

BGP flow specification version 7 

[See Supported Releases](#)

Results

The selected features are supported in following products/applications and releases:

| Product/Application | Supported Release(s) |        |        |        |        |        |        |        |        |        |        |        |        |         |         |         |         |        |  |
|---------------------|----------------------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|---------|---------|---------|---------|--------|--|
| MX5                 | Junos OS             |        |        |        |        |        |        |        |        |        |        |        |        |         |         |         |         |        |  |
|                     | 21.461               | 21.392 | 21.381 | 21.362 | 21.381 | 21.393 | 21.392 | 21.381 | 20.480 | 20.492 | 20.481 | 20.382 | 20.392 | 20.381  | 20.283  | 20.282  | 20.281  |        |  |
|                     | 20.183               | 20.182 | 20.181 | 19.483 | 19.482 | 19.481 | 19.382 | 19.382 | 19.381 | 19.283 | 19.282 | 19.281 | 19.183 | 19.182  | 19.181  | 18.483  | 18.482  |        |  |
|                     | 18.481               | 18.383 | 18.382 | 18.381 | 18.283 | 18.282 | 18.281 | 18.183 | 18.182 | 18.381 | 17.483 | 17.482 | 17.481 | 17.383  | 17.382  | 17.381  | 15.387  |        |  |
|                     | 15.186               | 15.177 | 15.185 | 15.184 | 15.184 | 15.183 | 15.185 | 15.184 | 15.183 | 15.184 | 15.183 | 15.182 | 15.181 | 12.3812 | 12.3811 | 12.3810 | 12.3809 | 12.388 |  |
| 12.387              | 12.386               | 12.385 | 12.384 | 12.383 | 12.382 | 12.381 |        |        |        |        |        |        |        |         |         |         |         |        |  |
| MX10                | Junos OS             |        |        |        |        |        |        |        |        |        |        |        |        |         |         |         |         |        |  |
|                     | 21.461               | 21.392 | 21.381 | 21.362 | 21.381 | 21.393 | 21.392 | 21.381 | 20.480 | 20.492 | 20.481 | 20.382 | 20.392 | 20.381  | 20.283  | 20.282  | 20.281  |        |  |
|                     | 20.183               | 20.182 | 20.181 | 19.483 | 19.482 | 19.481 | 19.382 | 19.382 | 19.381 | 19.283 | 19.282 | 19.281 | 19.183 | 19.182  | 19.181  | 18.483  | 18.482  |        |  |
|                     | 18.481               | 18.383 | 18.382 | 18.381 | 18.283 | 18.282 | 18.281 | 18.183 | 18.182 | 18.381 | 17.483 | 17.482 | 17.481 | 17.383  | 17.382  | 17.381  | 15.387  |        |  |
|                     | 15.186               | 15.177 | 15.185 | 15.184 | 15.184 | 15.183 | 15.185 | 15.184 | 15.183 | 15.184 | 15.183 | 15.182 | 15.181 | 12.3812 | 12.3811 | 12.3810 | 12.3809 | 12.388 |  |
| 12.387              | 12.386               | 12.385 | 12.384 | 12.383 | 12.382 | 12.381 |        |        |        |        |        |        |        |         |         |         |         |        |  |
| MX40                | Junos OS             |        |        |        |        |        |        |        |        |        |        |        |        |         |         |         |         |        |  |
|                     | 21.461               | 21.392 | 21.381 | 21.362 | 21.381 | 21.393 | 21.392 | 21.381 | 20.480 | 20.492 | 20.481 | 20.382 | 20.392 | 20.381  | 20.283  | 20.282  | 20.281  |        |  |
|                     | 20.183               | 20.182 | 20.181 | 19.483 | 19.482 | 19.481 | 19.382 | 19.382 | 19.381 | 19.283 | 19.282 | 19.281 | 19.183 | 19.182  | 19.181  | 18.483  | 18.482  |        |  |
|                     | 18.481               | 18.383 | 18.382 | 18.381 | 18.283 | 18.282 | 18.281 | 18.183 | 18.182 | 18.381 | 17.483 | 17.482 | 17.481 | 17.383  | 17.382  | 17.381  | 15.387  |        |  |
|                     | 15.186               | 15.177 | 15.185 | 15.184 | 15.184 | 15.183 | 15.185 | 15.184 | 15.183 | 15.184 | 15.183 | 15.182 | 15.181 | 12.3812 | 12.3811 | 12.3810 | 12.3809 | 12.388 |  |
| 12.387              | 12.386               | 12.385 | 12.384 | 12.383 | 12.382 | 12.381 |        |        |        |        |        |        |        |         |         |         |         |        |  |
| MX60                | Junos OS             |        |        |        |        |        |        |        |        |        |        |        |        |         |         |         |         |        |  |
|                     | 21.461               | 21.392 | 21.381 | 21.362 | 21.381 | 21.393 | 21.392 | 21.381 | 20.480 | 20.492 | 20.481 | 20.382 | 20.392 | 20.381  | 20.283  | 20.282  | 20.281  |        |  |
|                     | 20.183               | 20.182 | 20.181 | 19.483 | 19.482 | 19.481 | 19.382 | 19.382 | 19.381 | 19.283 | 19.282 | 19.281 | 19.183 | 19.182  | 19.181  | 18.483  | 18.482  |        |  |
|                     | 18.481               | 18.383 | 18.382 | 18.381 | 18.283 | 18.282 | 18.281 | 18.183 | 18.182 | 18.381 | 17.483 | 17.482 | 17.481 | 17.383  | 17.382  | 17.381  | 15.387  |        |  |
|                     | 15.186               | 15.177 | 15.185 | 15.184 | 15.184 | 15.183 | 15.185 | 15.184 | 15.183 | 15.184 | 15.183 | 15.182 | 15.181 | 12.3812 | 12.3811 | 12.3810 | 12.3809 | 12.388 |  |
| 12.387              | 12.386               | 12.385 | 12.384 | 12.383 | 12.382 | 12.381 |        |        |        |        |        |        |        |         |         |         |         |        |  |

# Support on Juniper, JunOS 7.3, August 2005?

## Router Vendors:

- Alcatel-Lucent SR OS 9.0R1
- Juniper JUNOS 7.3
- Cisco 5.2.0 for ASR and CRS [6]

Copyright © 2014 Juniper Networks, Inc.

# Support on Juniper, JunOS 7.2, May 2005!

## Flow Spec Status

IETF draft available at:

– <http://www.tcb.net/draft-marques-idr-flow-spec-03.txt>

- Implemented as of JunOS 7.2 (but not documented)
- At least three tier1/2 providers in process of production deployment
- Several security vendors announced integration
- Cisco complimentary TIDP proposal

8



<https://archive.nanog.org/meetings/nanog38/presentations/labovitz-bgp-flowspec.pdf>

# Support on Nokia, March 2011

Alcatel·Lucent



## 7750 SR OS Services Guide

Software Version: 7750 SR OS 9.0 r1  
March 2011  
Document Part Number: 93-0076-08-01



```
Entry      : fSpec-1-32767 - inserted by BGP FlowSpec
Description : (Not Specified)
Log Id     : n/a
Src. IP    : 0.0.0.0/0
Dest. IP   : 0.0.0.0/0
Protocol   : 6
ICMP Type  : Undefined
Fragment   : Off
Sampling   : Off
IP-Option  : 0/0
TCP-syn    : Off
Match action : Drop
Ing. Matches : 0 pkts
Egr. Matches : 0 pkts

Src. Port   : None
Dest. Port  : None
Dscp        : Undefined
ICMP Code   : Undefined
Option-present : Off
Int. Sampling : On
Multiple Option: Off
TCP-ack     : Off
```

```
Entry      : fSpec-1-49151 - inserted by BGP FlowSpec
Description : (Not Specified)
Log Id     : n/a
Src. IP    : 0.0.0.0/0
Dest. IP   : 0.0.0.0/0
Protocol   : 17
ICMP Type  : Undefined
Fragment   : Off
Sampling   : Off
IP-Option  : 0/0
TCP-syn    : Off
Match action : Drop
Ing. Matches : 0 pkts
Egr. Matches : 0 pkts

Src. Port   : None
Dest. Port  : None
Dscp        : Undefined
ICMP Code   : Undefined
Option-present : Off
Int. Sampling : On
Multiple Option: Off
TCP-ack     : Off
```

```
=====  
*A:Dut-C>config>filter#
```

# Support on Cisco, 2014

## Cisco Routers BGP FS Implementation



| Platform Hardware                                    | Support in Data Plane       |
|--|-----------------------------|
| ASR 9k – Typhoon LC (MOD80/160, 24-36x10G, 1-2x100G) | XR 5.2.0                    |
| ASR 9k – SIP700                                      | XR 5.2.2                    |
| ASR 9001(-S)   | XR 5.2.2                    |
| ASR 9k – Tomahawk (MOD200/400, 4-8-12x100G)          | XR 5.3.0                    |
| CRS-3 (Taiko) LC (1x100G, 14-20x10G, Flex)           | XR 5.2.0                    |
| CRS-X (Topaz) LC (4x100G, 40x10G, Flex)              | XR 5.3.2                    |
| NCS 6000   | XR 5.2.4 / 6.2.2 / roadmap* |
| XRv 9000   | 5.4.0 CP only / DP later    |
| NCS 5000 / NCS 5500                                  | In the roadmap              |
| ASR 1000   | IOS XE 3.15                 |
| CSR 1000v  | IOS XE 3.15                 |
| <b>NCS 5500 (Jericho+ w/ eTCAM)</b>                  | <b>XR 6.5.1</b>             |

Note: IOS XE introduced the support of BGP FS in 3.15 (but not as a controller role)

# Support on GoBGP, 2015

## IPv4/IPv6 FlowSpec

```
# Add a route
$ gobgp global rib -a {ipv4-flowsec|ipv6-flowsec} add match <MATCH> then <THEN>
<MATCH> : { destination <PREFIX> [<OFFSET>] |
  source <PREFIX> [<OFFSET>] |
  protocol <PROTOCOLS>... |
  fragment <FRAGMENTS>... |
  tcp-flags <TCP_FLAGS>... |
  port <ITEM>... |
  destination-port <ITEM>... |
  source-port <ITEM>... |
  icmp-type <ITEM>... |
  icmp-code <ITEM>... |
  packet-length <ITEM>... |
  dscp <ITEM>... |
  label <ITEM>... }...
<PROTOCOLS> : [&] [<|>] <PROTOCOL>
<PROTOCOL> : egp, gre, icmp, igmp, igp, ipip, ospf, pim, rsvp, sctp, tcp, udp, unknown, <DEC_NUM>
<FRAGMENTS> : [&] [!|] <FRAGMENT>
<FRAGMENT> : dont-fragment, is-fragment, first-fragment, last-fragment, not-a-fragment
<TCP_FLAGS> : [&] [!|] <TCP_FLAG>
<TCP_FLAG> : F, S, R, P, A, U, E, C
<ITEM> : [&] [<|>] <DEC_NUM>
<THEN> : { accept |
  discard |
  rate-limit <RATE> [as <AS>] |
  redirect <RT> |
  mark <DEC_NUM> |
  action [ sample | terminal | sample-terminal ] }...
<RT> : xxx:yyy, xxx.xxx.xxx.xxx:yyy, xxxxx:xxxxx:yyy, xxx.xxx:yyy

# Show routes
$ gobgp global rib -a {ipv4-flowsec|ipv6-flowsec}

# Delete route
$ gobgp global rib -a {ipv4-flowsec|ipv6-flowsec} del match <MATCH_EXPR>
```



# Support on Bird 2, 2017

## IPv4 Flowspec

`dst inet4`

Set a matching destination prefix (e.g. `dst 192.168.0.0/16`). Only this option is mandatory in IPv4 Flowspec.

`src inet4`

Set a matching source prefix (e.g. `src 10.0.0.0/8`).

`proto numbers-match`

Set a matching IP protocol numbers (e.g. `proto 6`).

`port numbers-match`

Set a matching source or destination TCP/UDP port numbers (e.g. `port 1..1023,1194,3306`).

`dport numbers-match`

Set a matching destination port numbers (e.g. `dport 49151`).

`sport numbers-match`

Set a matching source port numbers (e.g. `sport = 0`).

`icmp type numbers-match`

Set a matching type field number of an ICMP packet (e.g. `icmp type 3`).

`icmp code numbers-match`

Set a matching code field number of an ICMP packet (e.g. `icmp code 1`).

`tcp flags bitmask-match`

Set a matching bitmask for TCP header flags (aka control bits) (e.g. `tcp flags 0x03/0x0f`). The maximum length of mask is 12 bits (0xfff).

`length numbers-match`

Set a matching packet length (e.g. `length > 1500`).

`dscp numbers-match`

Set a matching DiffServ Code Point number (e.g. `dscp 8..15`).

`fragment fragmentation-type`

Set a matching type of packet fragmentation. Allowed fragmentation types are `dont_fragment`, `is_fragment`, `first_fragment`, `last_fragment` (e.g. `fragment is_fragment && !dont_fragment`).

# Support on Extreme, December 2018

## Overview

The focus of SLX-OS 18r.2.00 release is enhancing the Border Routing solution for SLX 9850, SLX 9540 as well as support for a new platform, the fixed form factor SLX 9640, for customers requiring larger route scale for border routing with Internet peering.

The following key software capabilities are added in this release:

- High IPv4, IPv6 route scale support on SLX 9640 to enable multiple full Internet peering tables on the same box using multiple VRFs
- Fast convergence at internet peering scale on bootup and peer, nexthop failures with BGP Prefix Independent Convergence(PIC).
- BGP Flowspec support for DDOS protection. This feature as described in RFC 5575 enables dissemination of filtering rules with standard BGP protocol to the border router (or from border router) so specific ACL filters can be applied to take various possible actions on DDOS attack traffic flows.
- BGP large community support per RFC 8092 to support 4-byte ASN in BGP communities attribute for policy handling.
- vSLX support for ESXi Hypervisor with vSLX install software 2.1.0

# Support on Arista, March 2020

## BGP Flowspec

The **EOS Release 4.21.3F** introduces support for BGP Flowspec, as defined in **RFC5575** and **RFC7674**. The typical use case is to filter or redirect DDoS traffic on edge routers.

BGP Flowspec rules are disseminated using a new BGP address family. The rules include both matching criteria used to match traffic, and actions to perform on the matching traffic. The rules are programmed into TCAM resources and applied on the ingress ports for which flowspec is enabled.

### Support for BGP flowspec + Release Updates

Written by Jason Shamberger | Posted on March 11, 2020 | Updated on February 22, 2021 | 2209 Views

EOS 4.21.3F introduces support for BGP Flowspec, as defined in RFC5575 and RFC7674. The typical use case is to filter

# 4.22.1 # 4.23.2F # 4.23.1 # Flowspec # 4.24.0 # 4.23.2 # 4.22.0

[Read More >](#)



# BGP Flow Spec challenges

- Limited number of BGP Flow Spec rules
- Lack of standard approach to retrieve packet and byte counters per rule
- Lack of proper rule validation
- Different hardware limitations
- Lack of interface to manage rules efficiently
- Weak integration with Netflow and IPFIX
- Lack of solid support for draft-ietf-idr-flowspec-redirect-ip-00

# BGP Flow Spec hardware limits: ASR 9000

## Cisco Bug: CSCuz29265 - [DOC]BGPFS dont-fragment and last-fragment match is not supported on A9k

### Last Modified

Sep 12, 2019

### Products (1)

Cisco ASR 9000 Series Aggregation Services Routers

### Known Affected Releases

5.2.4.FWDG 5.3.3.FWDG

### Description (partial)

#### Symptom:

dont-fragment and last-fragment match conditions are not supported by flowspec on the ASR9k (it's a HW limitation).

In the flowspec debug we will see following error:

```
RP/0/RP0/CPU0:Apr 12 10:31:37.458 : flowspec_mgr[1103]: %FLOWSPEC-3-MGR_CLASS_CREATE : Failed to create inline-class for flow Dest:1.0.0.2/32,Frag:=DF with actions Drop in table default:IPv4, overall:0x4081b400:'PBR' detected the 'warning' condition 'PBR PD': Not supported, 0x493bee30:'PBR' detected the 'warning' condition 'PBR PD': Not supported.
```

However it's not reflected in the documentation, for example:

[http://www.cisco.com/c/en/us/td/docs/routers/asr9000/software/asr9k\\_r5-2/routing/configuration/guide/b\\_routing\\_cg52xasr9k/b\\_routing\\_cg52xasr9k\\_chapter\\_011.html#task\\_16BCF875501E4C71812EC3188B318ABA](http://www.cisco.com/c/en/us/td/docs/routers/asr9000/software/asr9k_r5-2/routing/configuration/guide/b_routing_cg52xasr9k/b_routing_cg52xasr9k_chapter_011.html#task_16BCF875501E4C71812EC3188B318ABA)

#### Conditions:

match fragment-type dont-fragment  
or  
match fragment-type last-fragment

configured under flowspec class-map

# BGP Flow Spec hardware limits: Arista

- All matching components described in **RFC 5575** are supported, except for the following known caveats:
  - For TCP flags, the ECE, CWR, and NS flags are not supported.
  - For fragment flags, only the **Is a fragment (IsF)** bit is supported only for IPv4 packets. Combining source and destination ports and the Fragment flags in the same rule is not supported.

Similar to other TCAM features, the number of rules (BGP NLRI) that are supported in flowspec depend on the match criteria of each rule. Assuming that Flowspec is the only TCAM feature enabled on the switch, it attempts to use all of the TCAM space available (24K entries per chip) in the forwarding chip. Simple flowspec IPv4 rules will map to one entry, allowing a max of 24K rules. Simple IPv6 rules each take two entries, resulting in a max of 12K rules.

Some types of rules expand into multiple entries in the TCAM. Port ranges are a common example. Combining source and destination port ranges in a single rule multiplies the number of entries needed to cover all combinations, which can quickly consume all of the TCAM space.

The Flowspec and Flowspec Policer TCAM profiles support configuring the feature on up to seven VRFs starting with **EOS Release 4.24.1**. This scale can be adjusted with the number of bits in the feature's port qualifier size at the expense of removing other TCAM key fields.

Make-before-break policer allocation affects scaling limits.

# BGP Flow Spec and IPFIX, Netflow on Cisco

This Information Element describes the forwarding status of the flow and any attached reasons.

The layout of the encoding is as follows:

```
MSB - 0 1 2 3 4 5 6 7 - LSB
+-----+-----+
| Status| Reason code or flags |
+-----+-----+
```

See the Forwarding Status sub-registries at

[\[https://www.iana.org/assignments/ipfix/ipfix.xhtml#forwarding-status\]](https://www.iana.org/assignments/ipfix/ipfix.xhtml#forwarding-status).

Examples:

```
value : 0x40 = 64
binary: 01000000
decode: 01      -> Forward
        000000  -> No further information
```

```
value : 0x89 = 137
binary: 10001001
decode: 10      -> Drop
        001001  -> Bad TTL
```

## Forwarding Status (Value 89)

### Registration Procedure(s)

Expert Review

### Expert(s)

IE Doctors

### Reference

[\[RFC7270\]](#)

### Available Formats



CSV

| Value | Description | Reference                 |
|-------|-------------|---------------------------|
| 00b   | Unknown     | <a href="#">[RFC7270]</a> |
| 01b   | Forwarded   | <a href="#">[RFC7270]</a> |
| 10b   | Dropped     | <a href="#">[RFC7270]</a> |
| 11b   | Consumed    | <a href="#">[RFC7270]</a> |

Status 00b: Unknown



# FastNetMon: our community

- Site: <https://fastnetmon.com>
- GitHub: <https://github.com/pavel-odintsov/fastnetmon>
- Slack: <https://slack.fastnetmon.com/>
- Telegram: <https://t.me/fastnetmon>
- IRC: #fastnetmon at Libra Chat
- Discord: <https://discord.fastnetmon.com/>
- LinkedIn: <https://www.linkedin.com/company/fastnetmon/>
- Facebook: <https://www.facebook.com/fastnetmon/>
- Twitter: <https://twitter.com/fastnetmon>



# THANKS!

ANY QUESTIONS?

You can find me at:

- ◇ @odintsov\_pavel
- ◇ pavel@fastnetmon.com
- ◇ linkedin.com/in/podintsov

