BGP Flowspec

A look back...
... and a look forward

NANOG 85

Jeffrey Haas, D.E.
Juniper Networks
<jhaas@juniper.net>

© 2022 Juniper Networks

Q: What is automation?
A: Breaking the network at scale.

Attributed to Mark Tinka, NANOG mailing list, October 5, 2021

© 2022 Juniper Networks

What is BGP Flowspec?

It's a feature for distributing firewall rules into BGP.

The most common use case is to quickly deal with distributed denial of service (DDoS) attacks.

Specified in the IETF originally in RFC 5575. Updated in 2020 by RFC 8955.

- Original author list was a mix of router vendors, operators, and those working on DDoS detection and mitigation software.

© 2022 Juniper Networks

Whatever could go wrong…?

© 2022 Juniper Networks

Distributing firewall rules...

Firewalls are still the best place to deal with certain types of attacks.
It's also a good way to do clever things with your traffic!

They aren't constrained to destination-based forwarding.
(If RTBH, RFC 5635, can help you, do that instead!)

The problem is that *firewall* word.

© 2022 Juniper Networks

Distributing firewall rules...

The problem with firewalls

BGP Flowspec provides a limited set of well-known match operations, and a very small set of actions when you match. These things are well supported on most stateless firewalls.

Match Types	
Destination Prefix	ICMP Type
Source Prefix	ICMP Code
IP Protocol	TCP Flags
Port	Packet Length
Destination Port	DSCP (Diffserv Code Point)
Source Port	Fragment

Actions
Traffic Rate in Bytes (traffic-rate-bytes)
Traffic Rate in Packets (traffic-rate-packets)
Traffic-Action (traffic-action)
RT Redirect (rt-redirect)
Traffic Marking (traffic-marking)

© 2022 Juniper Networks

Distributing firewall rules...

The problem with firewalls

Except that anyone that has operated any large heterogenous firewall installation is wisely afraid of the phrase, “well-supported”:

- Are all of the operations actually supported? What if you’re missing a primitive in some part of your deployment?
- What if the primitives have unexpected limitations?
- How do the rules scale vs. their impact on packet forwarding rates?

- What about (whispers...) bugs?

© 2022 Juniper Networks

… into BGP

A flow specification rule is serialized into a byte-string in BGP's NLRI for matches, and BGP Extended Communities for actions.

- It's one of the few formats that serialize somewhat arbitrary things into NLRI.
  The other big one is BGP-Link State (LS).

This lets a Service Provider rapidly deploy firewall rules to the routers that can implement it. Route reflectors can make it easy!

© 2022 Juniper Networks

… into BGP

It lets a Service Provider accept flowspec rules from their *customers* that may need traffic mitigated by their upstream provider.

It provides the ability to validate those rules to ensure they're only enforced on destinations that customer controls.

© 2022 Juniper Networks

… into BGP

Serialization

- Various match components support lists.
  - Universal canonicalization isn’t practicable.
- The primary format isn’t the traditional Type-Length-Value (TLV) format, it’s Type-Value with the Length field inferred.
  - While the RFC 5575 authors intended NLRI with unknown match components to be “opaque”, not a single implementation treated it this way.
  - This mean incremental deployment of new features was compromised.

© 2022 Juniper Networks

… into BGP

Serialization

There’s enough places in the protocol where you can end up with improperly formed NLRI. RFC 5575 had a lot of missing edge cases in error handling. That was one of the largest motivations for the RFC 8955 update.

This meant you got session resets if you messed up.

© 2022 Juniper Networks

… into BGP

Rule complexity and ordering

To top it off, since a given router might be given Flowspec routes from many places, how do you arrange the firewall rules? Remember, in firewalls, *order is important*.

A complex bit of rule ordering. (RFC 8955, §5.1) Roughly:

- Sort by match component type first. The rule list for current features mostly makes sense for that!
- Sort by length. The more stuff you're matching, the "more specific" the route is.

© 2022 Juniper Networks

… into BGP

Rule complexity and ordering

This is good for ”simple” DDoS. But it’s not always great for the firewalls.

Currently, you can’t control explicit rule order.

© 2022 Juniper Networks

… into BGP

Rule complexity and ordering

The protocol supports far more flexibility for rulesets than many platforms can configure in their user interfaces.

- What about (whispers...) bugs?

© 2022 Juniper Networks

”BGP is a Swiss Army knife.
It doesn’t even have a handle, it’s all knives.”

Dr. Tony Przygienda

© 2022 Juniper Networks

Sometimes the best tool for the job will cut you

Many large-scale Service Providers successfully leverage BGP Flowspec as part of their DDoS mitigation toolset.

A vendor-neutral tool that distributed firewall rules in something besides BGP would run into similar firewall problems. For example, NETCONF, gRPC, etc.

Having the distribution in-line to the control plane has nice properties.

What can we do to help do the job better?

© 2022 Juniper Networks

BGP Flowspec v2

Solving the encoding problems

- New work in IETF's IDR Working Group.
- Make the NLRI format explicit Type-*Length*-Value fields.
  This means we can incrementally deploy new features safely without session resets.

© 2022 Juniper Networks

BGP Flowspec v2

Solving the encoding problems

- Explicit rule ordering.
  Permit the operator to not rely on protocol rule ordering when it doesn't suit their filtering needs.
  Also can address firewall optimizer issues.
- Smarter action clustering.
  Remove ambiguity in rule ordering.
  Provide framework for rule chaining.

© 2022 Juniper Networks

BGP Flowspec rule scoping

Work under discussion

Not all devices speaking Flowspec can support future features.

- Some don’t support the ones that are in the core of the RFC.

Limit the propagation of Flowspec rules only among devices that understand how they’re implemented.

- Can permit a route reflector deployment to selectively deploy rules.

© 2022 Juniper Networks

BGP Flowspec node capability discovery

How to discover platform limitations to enable smarter controllers?

There's a loose area of work in IETF that is discussing how to deal with feature and feature capability discovery:

- MPLS maximum segment depth for MPLS is an easy example.
- Some of these features are showing up in the IGPs
- There's a desire to keep the full set of these *outside* of the IGP.
  - There's tension on wanting them to be kept along-side the IGP so that tools can readily map capabilities to place in topology.
- For Flowspec, supported capabilities can help a controller build better rules.

© 2022 Juniper Networks

Work happening in IETF that needs Flowspec v2

- Flowspec v2 core Internet-Draft
- Flowspec Payload Match
- Flowspec for L2VPN
- Flowspec for Tunneled Traffic
- Flowspec for SRv6
- Flowspec with SRv6 policy
- Flowspec for APN
- Flowspec load balancing group community
- Detnet Flow Mapping

© 2022 Juniper Networks

Operator participation wanted

Flowspec v2 work will be happening in IETF's IDR Working Group.

If you have interest in the future of the protocol or other operational considerations for it, please comment (or contribute) on the work!

- https://www.ietf.org/mailman/listinfo/idr

Talk to me, and other vendors, about building safer Flowspec.

© 2022 Juniper Networks

Thank You

© 2022 Juniper Networks