BGP Path Attribute Filtering

IEPG - IETF 123, Madrid

Jeffrey Haas jhaas@juniper.net, John Scudder jgs@juniper.net

BGP Path Attribute History

- We've had Path Attributes and have had a transitive bit in BGP since BGP-2 (RFC 1163)
- Path Attributes are BGP's way of pairing tuples of route properties with network destinations (NLRI). Path Attributes have code points 0..255 (1 byte).
- The transitive bit's usage is:
  - If clear, then the attribute is non-transitive, and if the receiving implementation doesn't understand the attribute, it should discard that attribute.
  - If set, then if the attribute isn't understood, you mark it with the "partial" bit, but otherwise should just pass it on.
- ***Unrecognized transitive Path Attributes are how BGP incrementally deploys new features.***

What's "well known"

- RFC 4271 (BGP-4) has a minimal set of protocol path attributes defined that every implementation has to understand: ORIGIN, AS_PATH, LOCAL_PREF, MULTI_EXIT_DISC (MED), NEXT_HOP, ATOMIC_AGGREGATE, AGGREGATOR.
- There's a lot more: https://www.iana.org/assignments/bgp-parameters/bgp-parameters.xhtml#bgp-parameters-2
- Implementations tend to understand more than the basics: Route reflection, Communities, Multi-Protocol (IPv6, etc.), 4-byte AS_PATHs.
- All of these common things are ***optional.***

Some may call this junk, me I call them treasures.

- Many new features for BGP define additional Path Attributes.
- Over recent years, non-Internet use cases for BGP have created new Path Attributes where the Internet isn’t the main use case.
  - VPN features are a common example
  - BGP-LS
  - Some of these attributes change forwarding and route selection!
- But new Internet-focused features as well, such as Large BGP Communities!

… or just junk

- Much of the time, to ease incremental deployment, these new features were made transitive.
  - However, inconsistent care has been taken to deal with these new attributes leaving their appropriate scope.
  - This leads to “attribute escape”.
  - Sometimes this escape causes incorrect forwarding.
  - And often, escaped attributes are associated with crashes or security issues.
- Much of the original motivation for new RFC 7606 error handling procedures was to deal with “optional transitive nonsense”.

Dealing with other people's junk

- The issues associated with bugs or forwarding issues caused by new path attributes have led to implementations creating Path Attribute *filtering* features.
  - Some stop routes with specific attributes.
  - Some strip those attributes.
  - Some locally ignore those attributes but pass along routes with them.
  - ***These operations are not standardized.***
- ***Filtering Path Attributes breaks incremental deployment of new features!***
  - If you're a transit ISP, your filtering is making feature choices for your downstream customers.

Dealing with the tension on new features

- When a provider considers BGP Path Attribute filtering, they're making choices for themselves and their downstream customers.
- Like other security policies, consistent enforcement, awareness of feature use and efficacy, and agility to update policies is important.
- What's problematic is that you can't easily shop for ISPs that do or do not filter things. You usually find out because you're negatively impacted by such filtering.
- If you're a leaf AS, the story is easier to manage.

Handling this in the protocol

- There have been some prior efforts to talk about scoping Path Attributes generally, however they've not been successful.
  - Probably a thing that can only consistently happen in BGP-5.
- Awareness of the issue means that when we design new features, we can be mindful of scoping and escape considerations in that design.
  - (And even with such awareness, a recent feature changed the scope of where it was used to include the Internet. It became implicated in recent outages. We need to do better.)

Improving filtering

- Operators are filtering today. However, it's a silent feature, and there's no visibility when it's used.
- It'd be useful if providers could publish their policies at a peering session, and if both sides would help enforce them.
  - There's a proposal in IDR to discuss doing this: draft-haas-idr-path-attribute-filtering

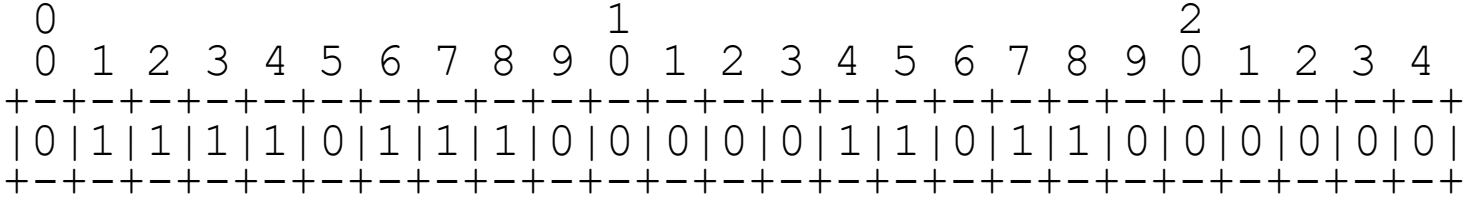
Path Attribute Filtering Capability -00

Capability Code of (TBD).

Capability Length of 3..32 octets.

Capability Value contains a **bit-string** where a bit is set if the underlying BGP Path Attribute is desired to be advertised by this BGP speaker to the remote BGP speaker.

Example encoding for Capability Value:



Origin (1),
AS_PATH (2),
NEXT_HOP (3),
MULTI_EXIT_DISCR (4),
ATOMIC_AGGREGATE (6),
AGGREGATOR (7),
COMMUNITIES (8),
MP_REACH_NLRI (14),
MP_UNREACH_NLRI (15),
AS4_PATH (17),
AS4_AGGREGATOR (18).

This encoding will look familiar to those who know the SNMP BITS type.

Default filtering policy

- The real conversational point for operators is… what's the default policy?
  - Do we permit unknown by default?
  - Do we filter it? If so, discard the attribute or block routes?
- These choices will change what the Internet looks like tomorrow.
  - (Even without our protocol feature…)

Let’s talk

- Filtering is already happening.
- This feature would help create per-router visibility.
- Should broader filtering policy be published somewhere by a provider?
- What additional operational visibility should these features have?
- Should transitive attributes be filtered by default, or not?
  - Filtering by default would be a seismic change to our past assumptions for new feature deployment.

Thanks