BGP Attribute Escape

draft-haas-idr-bgp-attribute-escape-00

IETF 117, San Francisco

Jeffrey Haas <jhaas@juniper.net>

What is BGP Attribute Escape?

- Any circumstance where a BGP Path Attribute attached to a route manages to be propagated outside of its intended scope is an "escape".
- While this word has a sense of a prisoner breaking out of a jail, often the sense is closer to a dog wandering off of its leash and running down the road.
  - Normal occurrences of this are not typically malicious events.
  - Sometimes this is a configuration issue, but sometimes no configuration can help.

Why was this document written?

- The BGP RFC specifications are good about describing how Path Attribute propagation works.
- Individual features are good about discussing how those features interact with their new Path Attributes.
- The intrinsic behavior of BGP when deploying new features means that there are unconsidered consequences impacting those attributes not staying “where they belong”.
- These issues should be readily documented so that they can be recognized, mitigated in their design and deployment of their features, and help future authors avoid these mistakes.

What's the problem with attribute escape?

- Routers in networks that receive BGP routes with features that have local significance and insufficient scoping information might try to use that information locally.
  - A prefix SID that gets put on a route might cause mis-routing or blackholes in your network.
  - A tunnel encapsulation attribute may try to tunnel traffic with the wrong encapsulation, or even try to send it to a remote network!
  - Some local signaling information like an extended community might cause a VPN feature to misbheave.
  - Route selection may be negatively impacted.

Optional Transitive Nonsense

- The motivation for our updated RFC 7606 error handling procedures was caused by attribute escape of attribute 128 (RFC 6368) causing networks that had newer versions of the feature to crash when receiving the older version of that feature, or vice-versa.
- While the error handling procedures helps us not-crash or at least avoid BGP session resets, the bigger issue is the attribute typically shouldn’t have gotten that far in the first place.

Recent example of the issue - Entropy Label

- The BGP Entropy label feature was originally published in RFC 6790. It had no scoping information in it.
- Oops! The Juniper authors realized that it may have escape, and RFC 7447 was quickly published to say “don’t use this!”
- Oops! The Juniper authors had released a non-RFC 6790 compatible feature in the same code point. (And published an after-the-fact Internet-Draft documenting the running code.)
- Oops! Other vendors implemented the original RFC 6790 feature in spite of the warnings in RFC 7447. Juniper routers receiving this would reset sessions if RFC 7606 treat-as-withdraw behaviors weren’t enabled.
- This was able to happen in most cases because the routes from those vendors had RFC 6790 escape.
- (Long term fix: draft-ietf-idr-next-hop-capability)

Recent example of the issue: DPATH

- https://datatracker.ietf.org/doc/html/draft-ietf-bess-evpn-ipvpn-interworking defines the DPATH attribute
- This attribute impacts route selection and is intended for EVPN and closely coupled L3VPN networks.
- If the attribute escapes, it can cause inconsistent route selection and potentially forwarding loops in networks that inconsistently deploy the feature.
- Provides a good example of family scoping considerations, and also the need for features impacting route selection to be in a consistently deployed domain.

This is really about scoping

- We start more formally discussing attribute scoping:
- AS-level scoping is very common.
  - We can mitigate most of these issues by including a “scoping AS” in new attributes.
- Some features are really scoped based on nexthops getting changed.
  - The BGP nexthop-capability referenced earlier is a way to enforce that.
- Consistent route selection issues rely on having the feature deployed consistently in any iBGP domain where the nexthops remain unchanged.
  - We don’t have a consistent mechanism for dealing with this yet!

Communities can escape too

- For extended communities, most VPN extended communities have transitive EC semantics.
  - They can leak information that might be sensitive.
  - Cleanup procedures, especially for routes leaking from/to VRF instances aren't consistent across implementations.
  - In some cases, they may have properties that impact traffic. E.g. link-bw.
- Even for regular and large communities, things go too far.
  - RFC 7454 discusses some of this.
  - Randy Bush has presented many nice papers at past IETFs on issues this can cause.

Not oops

- The issues described in the prior slides can, in some places, be maliciously exploited.
- This is another reason to make sure there’s consistent attention on the issues described in this document.

Mitigations

- Explicit filtering of Path Attributes is a feature present in several implementations.
- Indiscriminate use of such filtering can negatively impact the deployment of new BGP features.
- Implementations can use smarter filtering technologies.
  - More “easy mode” profiles for “scrubbing” routes.
  - Internet-Drafts for BGP features should have a mandatory section on this! It lives somewhere between Security and Operational Considerations.
  - Writing such profiles for existing features should be taken up by IDR.
- If we ever get around to BGP-5, or MP-UPDATE-v2, these considerations should be part of the core protocol.

Next steps

- Continue to socialize this issue where BGP is used.
    - Need presentations in BESS (perhaps next IETF).
    - Worth presenting at operational forums like NANOG?
- Ask IDR to adopt this as an active working group document and add to it.
    - Scrubbing profiles for existing features might go in this document or new ones. Compare vs. RFC 7606 work.
    - Mandate new sections in BGP Internet-Drafts to account for and mitigate these issues?
- Figure out the status and publish as a RFC.

Questions???