Workgroup: Inter-Domain Routing
Internet-Draft:
draft-haas-idr-bgp-attribute-escape-01
Published: 2 February 2024
Intended Status: Informational
Expires: 5 August 2024
Authors: J. Haas
Juniper Networks

BGP Attribute Escape

Abstract

BGP-4 [RFC 4271] has been very successful in being extended over the years it has been deployed. A significant part of that success is due to its ability to incrementally add new features to its Path Attributes when they are marked "optional transitive". Implementations that are ignorant of a feature for an unknown Path Attribute that are so marked will propagate BGP routes with such attributes.

Unfortunately, this blind propagation of unknown Path Attributes may happen for features that are intended to be used in a limited scope. When such Path Attributes inadvertently are carried beyond that scope, it can lead to things such as unintended disclosure of sensitive information, or cause improper routing. In their worst cases, such propagation may be for malformed Path Attributes and lead to BGP session resets or crashes.

This document calls such inadvertent propagation of BGP Path Attributes, "attribute escape". This document further describes some of the scenarios that leads to this behavior and makes recommendations on practices that may limit its impact.

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 5 August 2024.

Table of Contents

1. Introduction

1.1. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "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.

2. BGP Path Attributes and Transitivity

BGP-4 [RFC4271] carries routing information in UPDATE messages. BGP UPDATEs pair sets of network layer reachability information (NLRI), i.e. destinations, with properties for those destinations. Those properties are carried in Path Attributes (Section 4.3 of [RFC4271]). The BGP-4 RFC defines a base set of Path Attributes, several of which are "mandatory", and two that are optional.

Optional Path Attributes are flagged to be either "transitive" or "non-transitive". This flag controls how a BGP implementation that does not recognize a Path Attribute will handle it. In particular, unrecognized non-transitive Path Attributes, "MUST be quietly ignored and not passed along to other BGP peers". However, if the attribute is non-transitive but recognized, the procedures documented for that feature will control how it is propagated by such an implementation.

Emphasizing this point: Non-transitive Path Attributes are only guaranteed to be dropped during BGP route propagation by implementations that do not recognize them.

Unrecognized transitive Path Attributes, if they are accepted by the implementation (and they usually are), will be propagated. However, when they are unrecognized, the "Partial" flag is set indicating that at least one BGP speaker in the propagation path did not recognize the feature. In practice, this flag is solely informational and not used by implementations to change their processing of the attribute when they recognize it.

3. Motivation to Make New Path Attributes Transitive, and When It's Not Safe

Authors of new BGP features often desire to make their new Path Attributes to be transitive in order to avoid having to upgrade BGP speakers in their networks to support the new feature. A particular case where this is desired is to avoid upgrading BGP route reflectors [RFC4456]. When it's the case that a feature only needs to be supported by a partial set of BGP speakers in the network, this often works well.

A particular case where this strategy fails is when the feature alters BGP's Decision Process (Section 9.1 of [RFC4271]); i.e. route selection. In general, it is necessary for all BGP speakers in an Autonomous System (AS) to agree on how to do route selection. Inconsistent route selection may lead to routing or forwarding loops.

A consequence of this point is that any new feature that alters route selection needs to be consistently deployed within an AS. In most circumstances, this is true for the full scope of BGP route propagation where the next hop is unchanged. Thus, this applies to BGP Confederations [RFC5065] as well.

Non-transitive Path Attributes can be helpful for features that impact route selection. A consequence of their behavior is that once a route containing a non-transitive Path Attribute reaches a BGP speaker that is ignorant of its behavior, that attribute will be dropped. An example of this behavior is used in the AIGP [RFC7311] feature.

4. Attribute Escape

4.1. Attribute Scoping

Propagation scope is typically described as part of a BGP protocol extension's procedures. A common scoping boundary is intra-AS only; for example, drop at eBGP boundaries. BGP Capabilities [RFC5492] can also provide the ability to signal a scoping boundary. Another implicit scoping boundary are features that are AFI/SAFI specific ([RFC4760]); the scope is the contiguous domain of all devices that participate in that AFI/SAFI.

Non-transitive Path Attributes cannot be propagated outside of the scope of the set of BGP speakers that understand them. If such a Path Attribute is advertised "too far", it's a failure of the scoping mechanisms for that feature.

Transitive Path Attributes may similarly take advantage of the scoping mechanisms defined above, but only by devices that understand those Path Attributes.

4.2. Escape

Any circumstance where a BGP Path Attribute attached to a route manages to be propagated outside of its intended scope is an "escape".

A first example of escape are Path Attributes for new features that are intended to be dropped at eBGP boundaries. That is, the feature is intra-AS only. An example of such expected filtering is the BGP Tunnel Encapsulation Attribute (TEA) (Section 11 of [RFC9012]). BGP

speakers implementing this feature are expected to remove the TEA when sending routes to eBGP speakers by default. However, if a BGP speaker that is an eBGP router doesn't understand the TEA, it can't do this filtering. Similarly, the procedures recommend the TEA is removed by default at eBGP boundaries. In such cases, if the eBGP speaker doesn't understand the TEA, it won't be removed there either. When both failures happen, and if it's the case that an AS receives routes with the TEA from another AS the receiving AS may inadvertently use that TEA within their network.

A second example of escape is when attributes are expected to change at next hop reset boundaries. While this is typically the case for eBGP, this could be done elsewhere within a network by configuration. One example of this is the BGP Entropy Label feature [RFC6790]. A second example is the BGP Prefix-SID feature [RFC8669].

A third example of escape is when attributes leak across AFI/SAFIs. Layer-3 VPNs [RFC4364] operate by distributing routes learned from customer networks as IPv4 Unicast routes through a provider backbone using a different AFI/SAFI. Features that are used within the provider network, and attributes attached to those routes that are solely for use within the scope of the provider network and not intended for the customer network, need to be filtered when the route is again placed in a customer network context. However, it's often the case that filtering is not carefully done when leaking the routes back into the customer network (VRF). Unrecognized Path Attributes may be copied wholly. An example of such a provider-network-only feature is the ATTR_SET [RFC6368].

4.3. Community Escapes

BGP Communities of various forms ([RFC1997], [RFC4360], [RFC8092]) are often arbitrary operator-supplied route markup intended to have local or closely coordinated significance. RFC 1997 provides for a small set of well-known communities that have impact on BGP route scoping, and are reasonably well-respected by various implementations. RFC 1997 and RFC 8092 Large BGP communities often are distributed further than necessary, and in most circumstances simply contribute to "operational clutter".

However, some communities that are partially standardized but don't have consistent global significance may also impact traffic. An example of this is the BGP Blackhole Community [RFC7999].

The majority of RFC 4360 BGP Extended Communities in use tend to be for features related to VPN-feature signaling. Extended Communities have their own internal transitivity scoping, which is intended to be enforced at AS boundaries. However, such enforcement can only be implemented by BGP speakers that understand Extended Communities.

In general, while BGP community features are well-deployed and well understood, they may have similar escape issues. Section 11 of [RFC7454] offers some guidance on locally scrubbing such things.

4.4. Inadvertent Use

One of the main, and typically accidental impacts of attribute escape, are BGP speakers that have received such escaped attributes and make use of them. In some cases, these are features that require explicit configuration to be used. In many cases, BGP implementations will simply process received attributes and make use of them without additional configuration.

Such inadvertent use can have unexpected impacts on a network. Impacts may include unexpected route selection, mis-routed traffic, or traffic that blackholes.

Such behaviors may be maliciously exploited.

4.5. Outages Cause by Attribute Escape

While transitive Path Attributes have proved that BGP's extension mechanisms have done well over the years, they've also provided the vehicle for network incidents and outages. One form of this is when an implementation of a partially deployed feature receives a malformed transitive Path Attribute from a BGP speaker that didn't understand that attribute. The author has previously termed this problem, "optional transitive nonsense".

The underlying issue for such nonsense is the response of the receiving BGP speaker. Using core RFC 4271 procedures, such a malformed Path Attribute is reason to reset the BGP session. (Section 6.3 of [RFC4271]) However, since the BGP speaker that propagated the route may not have been the originator of that attribute, this penalizes the BGP speaker that propagated the route.

In far more unfortunate circumstances, such malformed attributes may exercise defects in the BGP stack and cause crashes. These crashes impacts the entire BGP speaker rather than individual sessions. Such issues may be maliciously exploited.

These issues were motivations for the "Revised Error Handling for BGP UPDATE Messages" document, [RFC7606].

5. Mitigating Attribute Escape

5.1. Explicit Permit Lists, and Their Dangers

One strategy to mitigate these sort of issues can be explicit permit lists. Rather than simply propagate all unknown Path Attributes,

lists of permitted Path Attributes by Attribute Type Code may be propagated through the network. However, this methodology has two strong negative impacts:

Similar to running any other large permit list infrastructure, as is often seen in large network firewall deployments, maintaining the lists themselves can be immensely burdensome at an operational level. Once a policy has been devised for the deployment, it must be consistent deployed to have good effect.

Additionally, this methodology is an "attack" on the feature that has provided for BGP's successful incremental deployment of new features. This struggle is a common one for users in firewalled networks where attempts to use new features are thwarted by the operators. However, in this case, this can impact deployment of desirable new BGP features for the Internet. Consider, for example, [RFC9234] which is intended to provide additional protection vs. BGP route leaking. Aggressive filtering may prevent such new features from being deployed.

Aggressive filtering may be more reasonable in some contexts. For example, a Layer-3 VPN customer may benefit from aggressive default Path Attribute filtering as it provides protection from attribute escape from their provider. However, in cases where newer BGP features are desired to be carried across the VPN, additional coordination with their service provider may be required.

5.2. Stronger Implementation Filtering Mechanisms

IETF review practices have evolved to start discussing attribute escape and its impacts. This review may lead to text recommending where filtering should be done and how. (See again the example for the Tunnel Encapsulation Attribute, Section 11 of [RFC9012].) It is important that implementors and vendors translate these filtering requirements into easy to configure mechanisms for enforcement.

Current IETF practice is to include an Operational Considerations ([RFC5706]) section in its documents. Such considerations are a hint (or requirement!) to implementors to create mechanisms that permit safe management of features. This is especially true during incremental or partial deployment of new features.

5.3. Scoping by Protocol Feature

Similar to the accepted truism that IETF considered security after the fact in its protocols, most older BGP features have not considered what to do about attribute escape. Understanding now that these are issues impacting incremental deployment of new features and their operations, it's possible to consider including scoping directly in new extensions.

5.3.1. AS Scoping

Many BGP features are intended to be scoped to one or more ASes. As described above, such mechanisms can't rely on all devices in an AS to have implemented a feature that's built using transitive Path Attributes.

New features that are intended to be AS-Scoped SHOULD consider including an scoping AS number, or AS number list, as part of the attribute. BGP speakers that use such features should have as part of their configuration a list of non-local ASes that they may wish to trust for utilizing such new features. This can permit a single "scope AS" to be included in the feature with a set of cooperating providers using different ASes able to enforce the scoping of such routes.

Implementations of such AS-Scoped features SHOULD have mechanisms that permit filtering of the feature's Path Attributes at AS boundaries. While the scoping-AS described above can help avoid issues with inadvertent usage, it can't help vs. intentional malicious exploitation where the scoping-AS is spoofed.

The BGP Wide Communities feature ([I-D.ietf-idr-wide-bgp-communities]) includes such an AS-Scoping mechanism.

5.3.2. Next Hop Scoping

Some BGP features rely on their semantics based on when the next hop of the route has been changed. In circumstances where the feature is only correct at such boundaries, embedding the "last set next hop" can be used to detect when a Path Attribute has propagated past a device that has reset the next hop without the associated change to the feature.

New features that are intended to next hop scope SHOULD consider including the next hop that was last present when the feature is being used.

An example of this methodology is the BGP Router Capabilities Attribute [I-D.ietf-idr-entropy-label].

5.3.3. Changes to the BGP protocol

More general changes to the BGP protocol could be considered to try to standardize some of these behaviors. One such approach has been documented, but not deployed in "Constrain Attribute announcement within BGP" [I-D.ietf-idr-bgp-attribute-announcement].

Similar approaches to this, or more protocol generalized forms of the other advice above, might be beneficial inputs if the core BGP specification is ever revised.

6. IANA Considerations

This memo includes no request to IANA.

7. Security Considerations

This document highlights properties of the BGP protocol and situations where its defined behavior for propagating Path Attributes may lead to inadvertent disclosure of information, improper routing, or even session resets and crashes. Such behaviors can be maliciously exploited.

8. References

8.1. Normative References

[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>.

[RFC4271] Rekhter, Y., Ed., Li, T., Ed., and S. Hares, Ed., "A Border Gateway Protocol 4 (BGP-4)", RFC 4271, DOI 10.17487/RFC4271, January 2006, <https://www.rfc-editor.org/info/rfc4271>.

8.2. Informative References

[RFC1997] Chandra, R., Traina, P., and T. Li, "BGP Communities Attribute", RFC 1997, DOI 10.17487/RFC1997, August 1996, <https://www.rfc-editor.org/info/rfc1997>.

[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>.

[RFC4360] Sangli, S., Tappan, D., and Y. Rekhter, "BGP Extended Communities Attribute", RFC 4360, DOI 10.17487/RFC4360, February 2006, <https://www.rfc-editor.org/info/rfc4360>.

[RFC4364] Rosen, E. and Y. Rekhter, "BGP/MPLS IP Virtual Private Networks (VPNs)", RFC 4364, DOI 10.17487/RFC4364, February 2006, <https://www.rfc-editor.org/info/rfc4364>.

[RFC4456] Bates, T., Chen, E., and R. Chandra, "BGP Route Reflection: An Alternative to Full Mesh Internal BGP

(IBGP)", RFC 4456, DOI 10.17487/RFC4456, April 2006, <https://www.rfc-editor.org/info/rfc4456>.

[RFC4760] Bates, T., Chandra, R., Katz, D., and Y. Rekhter, "Multiprotocol Extensions for BGP-4", RFC 4760, DOI 10.17487/RFC4760, January 2007, <https://www.rfc-editor.org/info/rfc4760>.

[RFC5065] Traina, P., McPherson, D., and J. Scudder, "Autonomous System Confederations for BGP", RFC 5065, DOI 10.17487/RFC5065, August 2007, <https://www.rfc-editor.org/info/rfc5065>.

[RFC5492] Scudder, J. and R. Chandra, "Capabilities Advertisement with BGP-4", RFC 5492, DOI 10.17487/RFC5492, February 2009, <https://www.rfc-editor.org/info/rfc5492>.

[RFC5706] Harrington, D., "Guidelines for Considering Operations and Management of New Protocols and Protocol Extensions", RFC 5706, DOI 10.17487/RFC5706, November 2009, <https://www.rfc-editor.org/info/rfc5706>.

[RFC6368] Marques, P., Raszuk, R., Patel, K., Kumaki, K., and T. Yamagata, "Internal BGP as the Provider/Customer Edge Protocol for BGP/MPLS IP Virtual Private Networks (VPNs)", RFC 6368, DOI 10.17487/RFC6368, September 2011, <https://www.rfc-editor.org/info/rfc6368>.

[RFC6790] Kompella, K., Drake, J., Amante, S., Henderickx, W., and L. Yong, "The Use of Entropy Labels in MPLS Forwarding", RFC 6790, DOI 10.17487/RFC6790, November 2012, <https://www.rfc-editor.org/info/rfc6790>.

[RFC7311] Mohapatra, P., Fernando, R., Rosen, E., and J. Uttaro, "The Accumulated IGP Metric Attribute for BGP", RFC 7311, DOI 10.17487/RFC7311, August 2014, <https://www.rfc-editor.org/info/rfc7311>.

[RFC7454] Durand, J., Pepelnjak, I., and G. Doering, "BGP Operations and Security", BCP 194, RFC 7454, DOI 10.17487/RFC7454, February 2015, <https://www.rfc-editor.org/info/rfc7454>.

[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>.

[RFC7999] King, T., Dietzel, C., Snijders, J., Doering, G., and G. Hankins, "BLACKHOLE Community", RFC 7999, DOI 10.17487/

RFC7999, October 2016, <https://www.rfc-editor.org/info/rfc7999>.

[RFC8092] Heitz, J., Ed., Snijders, J., Ed., Patel, K., Bagdonas, I., and N. Hilliard, "BGP Large Communities Attribute", RFC 8092, DOI 10.17487/RFC8092, February 2017, <https://www.rfc-editor.org/info/rfc8092>.

[RFC8669] Previdi, S., Filsfils, C., Lindem, A., Ed., Sreekantiah, A., and H. Gredler, "Segment Routing Prefix Segment Identifier Extensions for BGP", RFC 8669, DOI 10.17487/RFC8669, December 2019, <https://www.rfc-editor.org/info/rfc8669>.

[RFC9012] Patel, K., Van de Velde, G., Sangli, S., and J. Scudder, "The BGP Tunnel Encapsulation Attribute", RFC 9012, DOI 10.17487/RFC9012, April 2021, <https://www.rfc-editor.org/info/rfc9012>.

[RFC9234] Azimov, A., Bogomazov, E., Bush, R., Patel, K., and K. Sriram, "Route Leak Prevention and Detection Using Roles in UPDATE and OPEN Messages", RFC 9234, DOI 10.17487/RFC9234, May 2022, <https://www.rfc-editor.org/info/rfc9234>.

[I-D.ietf-idr-bgp-attribute-announcement]
Patel, K., Uttaro, J., Decraene, B., Henderickx, W., and J. Haas, "Constrain Attribute announcement within BGP", Work in Progress, Internet-Draft, draft-ietf-idr-bgp-attribute-announcement-03, 12 October 2022, <https://datatracker.ietf.org/doc/html/draft-ietf-idr-bgp-attribute-announcement-03>.

[I-D.ietf-idr-entropy-label]
Decraene, B., Scudder, J., Henderickx, W., Kompella, K., Mohanty, M., Uttaro, J., and B. Wen, "BGP Next Hop Dependent Capabilities Attribute", Work in Progress, Internet-Draft, draft-ietf-idr-entropy-label-13, 9 October 2023, <https://datatracker.ietf.org/doc/html/draft-ietf-idr-entropy-label-13>.

[I-D.ietf-idr-wide-bgp-communities]
Raszuk, R., Haas, J., Lange, A., Decraene, B., Amante, S., and P. Jakma, "BGP Community Container Attribute", Work in Progress, Internet-Draft, draft-ietf-idr-wide-bgp-communities-11, 9 March 2023, <https://datatracker.ietf.org/doc/html/draft-ietf-idr-wide-bgp-communities-11>.

Acknowledgements

Thanks to Greg Skinner for suggested edits on the document.

Contributors

TBD

Author's Address

Jeffrey Haas
Juniper Networks
1133 Innovation Way
Sunnyvale, CA 94089
United States of America

Email: jhaas@juniper.net