Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: January 9, 2020

A. Azimov
E. Bogomazov
Qrator Labs
R. Bush
Internet Initiative Japan & Arrcus
K. Patel
Arrcus, Inc.
K. Sriram
US NIST

July 8, 2019

Route Leak Prevention using Roles in Update and Open messages
draft-ietf-idr-bgp-open-policy-06

Abstract

Route Leaks are the propagation of BGP prefixes which violate assumptions of BGP topology relationships; e.g. passing a route learned from one peer to another peer or to a transit provider, passing a route learned from one transit provider to another transit provider or to a peer. Today, approaches to leak prevention rely on marking routes by operator configuration, with no check that the configuration corresponds to that of the BGP neighbor, or enforcement that the two BGP speakers agree on the relationship. This document enhances BGP OPEN to establish agreement of the (peer, customer, provider, Route Server, Route Server client) relationship of two neighboring BGP speakers to enforce appropriate configuration on both sides. Propagated routes are then marked with an OTC attribute according to the agreed relationship, allowing both prevention and detection of route leaks.

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.

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 January 9, 2020.

Table of Contents

1. Introduction

BGP route leaks are BGP route(s) which were learned from transit provider or peer and then announced to another provider or peer. See [RFC7908]. These are usually the result of misconfigured or absent BGP route filtering or lack of coordination between two BGP speakers.

The mechanism proposed in [I-D.ietf-idr-route-leak-detection-mitigation] uses large-communities to attempt detection of route leaks. While signaling using communities is easy to implement, ut relies on operator maintained policy configuration which is too easily, and too often, misconfigured. Another problem may occur if the community signal is stripped, accidentally or maliciously.

This document provides configuration automation using 'BGP roles', which are negotiated using a new BGP Capability Code in OPEN message, [RFC5492] Sec 4. Either or both BGP speakers MAY be configured to require that this capability be agreed for the BGP OPEN to succeed.

A new BGP Path Attribute is specified that SHOULD be automatically configured using BGP roles. This attribute prevents networks from creating leaks, and detects leaks created by third-parties.

2. Peering Relationships

Despite uses of words such as "Customer," "Peer." etc.; these are not business relationships, who pays whom, etc. These are common terms to represent restrictions on BGP route propagation, sometimes known as the Gao-Rexford model [cite].

A Provider: MAY send to a customer all available prefixes.

A Customer: MAY send to a provider their own prefixes and prefixes learned from any of their customers. A customer MUST NOT send to a provider prefixes learned from its peers, from other providers, or from Route Servers.

A Route Server (RS) MAY send to a RS Client all available prefixes.

A Route Server Client (RS-client) MAY send to an RS its own prefixes and prefixes learned from its customers. A RS-client MUST NOT send to an RS prefixes learned from peers, from its providers, or from other RS(s).

A Peer: MAY send to a peer its own prefixes and prefixes learned from its customers. A peer MUST NOT send to a peer prefixes learned from other peers, from its providers, or from RS(s).

Of course, any BGP speaker may apply policy to reduce what is announced, and a recipient may apply policy to reduce the set of routes they accept. Violation of the above rules may result in route leaks so MUST not be allowed. Automatic enforment of these rules should significantly reduce configuration mistakes. While these enforcing the above rules will address most BGP peering scenarios, their configuration isn't part of BGP itself; therefore requiring configuration of ingress and egress prefix filters is still strongly advised.

3. BGP Role

BGP Role is new configuration option that SHOULD be configured on each BGP session. It reflects the real-world agreement between two BGP speakers about their relationship.

Allowed Role values for eBGP sessions are:

- Provider - sender is a transit provider to neighbor;
- Customer - sender is transit customer of neighbor;
- RS - sender is a Route Server, usually at internet exchange point (IX)
- RS-Client - sender is client of RS
- Peer - sender and neighbor are peers;

Since BGP Role reflects the relationship between two BGP speakers, it could also be used for more than route leak mitigation.

4. Role capability

The TLV (type, length, value) of the BGP Role capability are:

- Type - <TBD1>;
- Length - 1 (octet);
- Value - integer corresponding to speaker' BGP Role.

Value	Role name
0	Sender is Provider
1	Sender is RS
2	Sender is RS-Client
3	Sender is Customer
4	Sender is Peer

Table 1: Predefined BGP Role Values

5. Role correctness

Section 3 described how BGP Role encodes the relationship between two BGP speakers. But the mere presence of BGP Role doesn't automatically guarantee role agreement between two BGP peers.

To enforce correctness, the BGP Role check is used with a set of constrains on how speakers' BGP Roles MUST correspond. Of course, each speaker MUST announce and accept the BGP Role capability in the BGP OPEN message exchange.

If a speaker receives a BGP Role capability, it MUST check the value of the received capability with its own BGP Role (if it is set). The allowed pairings are (first a sender's Role, second the receiver's Role):

Sender Role	Receiver Role
Provider	Customer
Customer	Provider
RS	RS-Client
RS-Client	RS
Peer	Peer

Table 2: Allowed Role Capabilities

If the Role pair is not in the above table, a speaker MUST send a Role Mismatch Notification (code 2, sub-code <TBD2>).

5.1. Strict mode

A new BGP configuration option "strict mode" is defined with values of true or false. If set to true, then the speaker MUST refuse to establish a BGP session with a neighbor which does not announce the

BGP Role capability in the OPEN message. If a speaker rejects a connection, it MUST send a Connection Rejected Notification [RFC4486] (Notification with error code 6, subcode 5). By default, strict mode SHOULD be set to false for backward compatibility with BGP speakers that do not yet support this mechanism.

6. BGP Only To Customer attribute

The Only To Customer (OTC) BGP Attribute is a new optional, transitive BGP Path attribute with the Type Code <TBD3>.

This four byte attribute MUST apply the following policy:

1. If a route with OTC attribute is received from Customer or RS-client - it's a route leak and MUST be rejected.

2. If a route with OTC attribute is received from Peer and its value isn't equal to the neighbor's ASN - it's a route leak and MUST be rejected.

3. If a route is received from a Provider, Peer or RS and the OTC attribute has not been set it MUST be added with value equal to AS number of the neighbor (sender).

The egress policy MUST be:

1. A route with the OTC attribute set MUST NOT be sent to providers, peers, or RS(s).

2. If route is sent to customer or peer and the OTC attribute is not set it MUST be added with value equal to AS number of the sender.

Once the OTC attribute has been set, it MUST be preserved unchanged.

7. Enforcement

Having the relationship unequivocally agreed between the two peers in BGP OPEN is critical; the BGP implementations enforce the relationship irrespective of operator policy configuration errors.

Similarly, the application of that relationship on prefix propagation using OTC MUST BE enforced by the BGP implementations, and not exposed to user mis-configuration.

As opposed to communities, BGP attributes may not be generally modified or filtered by the operator. The router(s) enforce them. This is the desired property for the OTC marking. Hence, this document specifies OTC as an attribute.

8. Additional Considerations

As the BGP Role reflects the peering relationship between neighbors, it might have other uses. For example, BGP Role might affect route priority, or be used to distinguish borders of a network if a network consists of multiple ASs. Though such uses may be worthwhile, they are not the goal of this document. Note that such uses would require local policy control.

As BGP role configuration results in automatic creation of inbound/ outbound filters, existence of roles should be treated as existence of Import and Export policy. [RFC8212]

There are peering relationships which are 'complex'; e.g. when both parties are intentionally sending prefixes received from each other to their peers and/or upstreams. If multiple BGP peerings can segregate the 'complex' parts of the relationship, the complex peering roles can be segregated into different BGP sessions, and normal BGP Roles MUST be used on the non-complex sessions. No Roles SHOULD be configured on 'complex' BGP sessions, and OTC MUST be set by configuration on a per-prefix basis. There can be no measures to check correctness of OTC use if Role is not configured.

9. IANA Considerations

This document defines a new Capability Codes option [to be removed upon publication: http://www.iana.org/assignments/capability-codes/capability-codes.xhtml] [RFC5492], named "BGP Role", assigned value <TBD1> . The length of this capability is 1.

The BGP Role capability includes a Value field, for which IANA is requested to create and maintain a new sub-registry called "BGP Role Value". Assignments consist of Value and corresponding Role name. Initially this registry is to be populated with the data in Table 1. Future assignments may be made by a standard action procedure[RFC5226].

This document defines new subcode, "Role Mismatch", assigned value <TBD2> in the OPEN Message Error subcodes registry [to be removed upon publication: http://www.iana.org/assignments/bgp-parameters/bgp-parameters.xhtml#bgp-parameters-6] [RFC4271].

This document defines a new optional, transitive BGP Path Attributes option, named "Only To Customer", assigned value <TBD3> [To be removed upon publication: http://www.iana.org/assignments/bgp-parameters/bgp-parameters.xhtml#bgp-parameters-2] [RFC4271]. The length of this attribute is 0.

10. Security Considerations

This document proposes a mechanism for prevention of route leaks that are the result of BGP policy mis-configuration.

Deliberate sending of a known conflicting BGP Role could be used to sabotage a BGP connection. This is easily detectable.

A misconfiguration in OTC setup may affect prefix propagation. But the automation that is provided by BGP roles should make such misconfiguration unlikely.

11. Acknowledgments

The authors wish to thank Douglas Montgomery, Brian Dickson, Andrei Robachevsky, and Daniel Ginsburg for their contributions to a variant of this work.

12. References

12.1. Normative References

[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>.

[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>.

[RFC4486] Chen, E. and V. Gillet, "Subcodes for BGP Cease Notification Message", RFC 4486, DOI 10.17487/RFC4486, April 2006, <https://www.rfc-editor.org/info/rfc4486>.

[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>.

[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>.

12.2. Informative References

[I-D.ietf-idr-route-leak-detection-mitigation]
Sriram, K. and A. Azimov, "Methods for Detection and Mitigation of BGP Route Leaks", draft-ietf-idr-route-leak-detection-mitigation-10 (work in progress), October 2018.

[RFC5226] Narten, T. and H. Alvestrand, "Guidelines for Writing an IANA Considerations Section in RFCs", RFC 5226, DOI 10.17487/RFC5226, May 2008, <https://www.rfc-editor.org/info/rfc5226>.

[RFC7908] Sriram, K., Montgomery, D., McPherson, D., Osterweil, E., and B. Dickson, "Problem Definition and Classification of BGP Route Leaks", RFC 7908, DOI 10.17487/RFC7908, June 2016, <https://www.rfc-editor.org/info/rfc7908>.

[RFC8212] Mauch, J., Snijders, J., and G. Hankins, "Default External BGP (EBGP) Route Propagation Behavior without Policies", RFC 8212, DOI 10.17487/RFC8212, July 2017, <https://www.rfc-editor.org/info/rfc8212>.

Authors' Addresses

Alexander Azimov
Qrator Labs

Email: a.e.azimov@gmail.com

Eugene Bogomazov
Qrator Labs

Email: eb@qrator.net

Randy Bush
Internet Initiative Japan & Arrcus

Email: randy@psg.com

Keyur Patel
Arrcus, Inc.

Email: keyur@arrcus.com

Kotikalapudi Sriram
US NIST

Email: ksriram@nist.gov