Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: May 18, 2017

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

November 14, 2016

Route Leak Detection and Filtering using Roles in Update and Open messages
draft-ymbk-idr-bgp-open-policy-02

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 according to operator configuration options without any 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, internal) relationship of two neighboring BGP speakers to enforce appropriate configuration on both sides. Propagated routes are then marked with a eOTC and iOTC attributes according to agreed relationship allowing prevention and detection of route leaks.

Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" are to be interpreted as described in RFC 2119 [RFC2119] only when they appear in all upper case. They may also appear in lower or mixed case as English words, without normative meaning.

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 http://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 May 18, 2017.

Copyright Notice

Copyright (c) 2016 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (http://trustee.ietf.org/license-info) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction

For the purpose of this document, BGP route leaks are when a BGP route was learned from transit provider or peer and is announced to another provider or peer. See [I-D.ietf-grow-route-leak-problem-definition]. These are usually the result of misconfigured or absent BGP route filtering or lack of coordination between two BGP speakers.

[I-D.ietf-idr-route-leak-detection-mitigation] describes a method of marking and detecting leaks which relies on operator maintained markings. Unfortunately, in most cases, a leaking router will likely also be misconfigured to mark incorrectly. The mechanism proposed in that draft provides the opportunity to detect route leaks made by third parties but provides no support to strongly prevent route leak creation. The leak prevention still relies on communities which are optional and often missed due to mistakes or misunderstanding of the BGP configuration process.

It has been suggested to use white list filtering, relying on knowing the prefixes in the peer's customer cone as import filtering, in order to detect route leaks. Unfortunately, a large number of incidents in medium transit operators use a single prefix list as only the ACL for export filtering, without community tagging and without paying attention to the source of a learned route. So, if they learn a customer's route from their provider or peer - they will announce it in all directions, including other providers or peers. This misconfiguration affects a limited number of prefixes; but such route leaks will obviously bypass customer cone import filtering made by upper level upstream providers.

Also, route tagging which relies on operator maintained policy configuration is too easily and too often misconfigured.

This document specifies a new BGP Capability Code, [RFC5492] Sec 4, which two BGP speakers MAY use to ensure that they MUST agree on their relationship; i.e. customer and provider or peers. Either or both may optionally be configured to require that this option be exchanged for the BGP Open to succeed.

Also this document specifies a way to mark routes according to BGP Roles established in Open and a way to create double-boundary filters for prevention and detection of route leaks via a two new BGP Path Attributes.

2. Role Definitions

As many of these terms are used differently in various contexts, it is worth being explicit.

A Provider: sends their own routes and (possibly) a subset of routes learned from their other customers, peers, and transit providers to their customer.

A Customer: accepts 'transit routes' from its provider(s) and announces their own routes and the routes they have learned from the transitive closure of their customers (AKA their 'customer cone') to their provider(s).

A Peer: announces their routes and the routes from their customer cone to other Peers.

An Internal BGP Neighbor has one of the above relationships to another internal BGP AS.

A Complex BGP relationship is an attempt to allow those whose policy may vary by prefix. It is aptly named and the authors question its real utility.

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.

3. BGP Role

BGP Role is new mandatory configuration option which must be set per each address family. It reflects the real-world agreement between two BGP speakers about their business relationship.

Allowed Role values are:

- Provider - sender is a transit provider to neighbor;
- Customer - sender is customer of neighbor;
- Peer - sender and neighbor are peers;
- Internal - sender is part of an internal AS of an organization which has multiple ASs, or is a confederation, etc.
- Complex - sender has a non-standard relationship and wants to use manual per-prefix based role policies.

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	Undefined
1	Sender is Peer
2	Sender is Provider
3	Sender is Customer
4	Sender is Internal
5	Sender is Complex

Table 1: Predefined BGP Role Values

5. Role correctness

Section 3 described how BGP Role is a reflection of 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 corresponded. 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 SHOULD check value of the received capability with its own BGP Role. The allowed pairings are (first a sender's Role, second the receiver's Role):

Sender Role	Receiver Role
Peer	Peer
Provider	Customer
Customer	Provider
Internal	Internal
Complex	Complex

Table 2: Allowed Role Capabilities

In all other cases 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 peers which do not announce the BGP Role capability in their 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. Restrictions on the Complex role

The Complex role should be set only if the relationship between BGP neighbors can not be described using simple Customer/Provider/Peer roles. For a example, if neighbor is literal peer, but for some prefixes it provides full transit; the complex role SHOULD be set on both sides. In this case roles Customer/Provider/Peer should be set on per-prefix basis, keeping the abstraction from detection and filtering mechanisms (Section 7 and Section 8).

If role is not Complex all per-prefix role settings MUST be ignored.

7. BGP Internal Only To Customer attribute

The Internal Only To Customer (iOTC) attribute is a new optional, non-transitive BGP Path attribute with the Type Code <TBD3>. This attribute has zero length as it is used only as a flag.

There are four rules for setting the iOTC attribute:

1. The iOTC attribute MUST be added to all incoming routes if the receiver's Role is Customer or Peer;

2. The iOTC attribute MUST be added to all incoming routes if the receiver's Role is Complex and the prefix Role is Customer or Peer;

3. Routes with the iOTC attribute set MUST NOT be announced by a sender whose Role is Customer or Peer;

4. Routes with the iOTC attribute set MUST NOT be announced if by a sender whose Role is Complex and the prefix Role is Customer or Peer;

These four rules provide mechanism that strongly prevents route leak creation by an AS.

8. BGP External Only To Customer attribute

The External Only To Customer (eOTC) attribute is a new optional, transitive BGP Path attribute with the Type Code <TBD4>. This attribute is four bytes and contains an AS number of the AS that added the attribute to the route.

There are four rules for setting the eOTC attribute:

1. If eOTC is not set and the sender's Role is Provider or Peer, the eOTC attribute MUST be added with value equal to the sender's AS number

2. If eOTC is not set and the sender's Role is Complex and the prefix role is Provider or Peer, the eOTC attribute MUST be added with value equal to to the sender's AS number.

3. If eOTC is set, the receiver's Role is Provider or Peer, and its value is not the neighbor's AS number then the incoming route is route leak and MUST be given a lower local preference, or MAY be dropped.

4. If eOTC is set, the receiver's Role is Complex, the prefix role Role is Provider or Peer, and the eOTC value is not equal to the neighbor's AS number, then the incoming route is a route leak and MUST be given a lower local preference, or they MAY be dropped.

These four rules provide mechanism for route leak detection that is created by an distant party in the AS_Path.

9. Compatibility with BGPsec

For BGPsec [I-D.ietf-sidr-bgpsec-protocol] enabled routers, the Flags field will have a bit added to indicate that an eOTC attribute exists. The eOTC value will be automatically carried in AS field of the added Secure_Path Segment.

When a route is translated from a BGPsec enabled router to a non-BGPsec router, in addition to AS_PATH reconstruction, reconstruction MUST be performed for the eOTC attribute.. If Flag bit was set in one of Secure_Path Segments, the eOTC attribute SHOULD be added with the AS number of the segment in which it appears for the first time.

10. Additional Considerations

As the BGP Role reflects the relationship between neighbors, it can also have other uses. As an example, BGP Role might affect route priority, or be used to distinguish borders of a network if a network consists of multiple AS.

Though such uses may be worthwhile, they are not the goal of this document. Note that such uses would require local policy control.

This document doesn't provide any security measures to check correctness of per-prefix roles, so the Complex role should be used with great caution. It is as dangerous as current BGP peering.

11. 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, non-transitive BGP Path Attributes option, named "Internal 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.

This document defines a new optional, transitive BGP Path Attributes option, named "External Only To Customer", assigned value <TBD4> [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 4.

12. Security Considerations

This document proposes a mechanism for prevention and detection of route leaks that are the result of BGP policy misconfiguration. This includes preventing route leaks created inside an AS (company), and route leak detection if a route was leaked by third party.

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

Deliberate mis-marking of the eOTC flag could be used to affect the BGP decision process, but could not sabotage a route's propagation.

BGP Role is disclosed only to an immediate BGP neighbor, so it will not itself reveal any sensitive information to third parties.

On the other hand, eOTC is a transitive BGP AS_PATH attribute which reveals a bit about a BGP speaker's business relationship. It will give a strong hint that some link isn't customer to provider, but will not help to distinguish if it is provider to customer or peer to peer. If eOTC is BGPsec signed, it can not be removed for business confidentiality.

13. Acknowledgments

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

14. References

14.1. 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, <http://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, <http://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, <http://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, <http://www.rfc-editor.org/info/rfc5492>.

14.2. Informative References

[I-D.ietf-grow-route-leak-problem-definition] Sriram, K., Montgomery, D., McPherson, D., Osterweil, E., and B. Dickson, "Problem Definition and Classification of BGP Route Leaks", draft-ietf-grow-route-leak-problem-definition-06 (work in progress), May 2016.

[I-D.ietf-idr-route-leak-detection-mitigation] Sriram, K., Montgomery, D., Dickson, B., Patel, K., and A. Robachevsky, "Methods for Detection and Mitigation of BGP Route Leaks", draft-ietf-idr-route-leak-detection-mitigation-03 (work in progress), May 2016.

[I-D.ietf-sidr-bgpsec-protocol] Lepinski, M. and K. Sriram, "BGPsec Protocol Specification", draft-ietf-sidr-bgpsec-protocol-15 (work in progress), March 2016.

[RFC5226] Narten, T. and H. Alvestrand, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 5226, DOI 10.17487/RFC5226, May 2008, <http://www.rfc-editor.org/info/rfc5226>.

Authors' Addresses

Alexander Azimov
Qrator Labs

Email: aa@qrator.net

Eugene Bogomazov
Qrator Labs

Email: eb@qrator.net

Randy Bush
Internet Initiative Japan

Email: randy@psg.com

Keyur Patel
Arrcus, Inc.

Email: keyurpat@yahoo.com

Kotikalapudi Sriram
US NIST

Email: ksriram@nist.gov