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, 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, RS, RS-client, internal) relationship of two neighboring BGP speakers to enforce appropriate configuration on both sides. Propagated routes are then marked with an iOTC attribute according to agreed relationship allowing prevention of route leaks.

Abstract

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 https://datatracker.ietf.org/drafts/current/.
1. Introduction

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 message and a way to create double-boundary filters for prevention of route leaks using the new BGP Path Attribute.

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[RFC7908]. 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] The mechanism proposed in that draft provides the opportunity to detect route leaks made by third parties but provides no mechanism to strongly prevent route leak creation.

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

2. Peering Relationships

Despite uses of words such as "Customer," "Peer." etc. described above 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.

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 peers, other providers, or RS.

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, providers, or other RS.

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, providers, or RS.

An Internal: MAY send all available prefixes through internal link.
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. But violation of rules marked MUST NOT may result in route leaks. While these peering relations cover 99% of possible scenarios, their configuration isn’t part of the BGP itself, thus requiring configuration of communities and corresponding egress prefix filters. The automation of this process may significantly decrease number of configuration mistakes.

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 peering relationship.

Allowed Role values for eBGP sessions are:

- Provider - sender is a transit provider to neighbor;
- Customer - sender is customer of neighbor;
- RS - sender is route server at internet exchange point (IX)
- RS-client - sender is client of RS at internet exchange point (IX)
- Peer - sender and neighbor are peers;
- Internal - sender and neighbor are part of the same organization.

For iBGP sessions, only the Internal role MAY be configured.

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.
Table 1: Predefined BGP Role Values

<table>
<thead>
<tr>
<th>Value</th>
<th>Role name</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Sender is Internal</td>
</tr>
<tr>
<td>1</td>
<td>Sender is Provider</td>
</tr>
<tr>
<td>2</td>
<td>Sender is RS</td>
</tr>
<tr>
<td>3</td>
<td>Sender is RS-Client</td>
</tr>
<tr>
<td>4</td>
<td>Sender is Customer</td>
</tr>
<tr>
<td>5</td>
<td>Sender is Peer</td>
</tr>
</tbody>
</table>

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 constraints 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):

<table>
<thead>
<tr>
<th>Sender Role</th>
<th>Receiver Role</th>
</tr>
</thead>
<tbody>
<tr>
<td>Internal</td>
<td>Internal</td>
</tr>
<tr>
<td>Provider</td>
<td>Customer</td>
</tr>
<tr>
<td>Customer</td>
<td>Provider</td>
</tr>
<tr>
<td>RS</td>
<td>RS-Client</td>
</tr>
<tr>
<td>RS-Client</td>
<td>RS</td>
</tr>
<tr>
<td>Peer</td>
<td>Peer</td>
</tr>
</tbody>
</table>

Table 2: Allowed Role Capabilities

In case of any other pairs of roles, 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 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. 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 of iOTC attribute usage:

1. The iOTC attribute MUST be added to all incoming routes if the receiver’s Role is Customer, Peer, or RS-client;

2. Routes with the iOTC attribute set MUST NOT be announced by a sender whose Role is Customer, Peer, or RS-client;

3. A sender MUST NOT include iOTC in UPDATE messages advertised to eBGP neighbor if its Role isn’t Internal.

4. If iOTC is contained in an UPDATE message from eBGP speaker and receiver’s Role isn’t Internal then this attribute MUST be removed.

These rules provide mechanism to strongly prevent route leak creation by an AS.

7. Attribute or Community

Having the relationship hard set by agreement between the two peers in BGP OPEN is critical; the routers enforce the relationship irrespective of operator policy configuration errors.

Similarly, it is critical that the application of that relationship on prefix propagation using iOTC is enforced by the router(s), and minimally exposed to user mis-configuration. There is a question whether the iOTC marking should be an attribute or a well-known community.
There is a long and sordid history of mis-configurations inserting incorrect communities, deleting communities, ignoring well-known community markings etc. In this mechanism’s case, an operator could, for example, accidentally strip the well-known community on receipt.

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 iOTC marking. Hence, this document specifies iOTC as an attribute.

8. Compatibility with BGPsec

As the iOTC attribute is non-transitive, it is not seen by or signed by BGPsec [RFC8205].

9. Additional Considerations

As the BGP Role reflects the peering 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 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]

This document doesn’t provide any security measures to check correctness of iOTC usage if role isn’t configured.

10. 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
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.

11. 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.

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

12. Acknowledgments

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

13. References

13.1. Normative References


13.2. Informative References

[I-D.ietf-idr-route-leak-detection-mitigation]


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
Email: randy@psg.com