IDR and SIDR K. Sriram, Ed.
Internet-Draft USA NIST
Intended status: Standards Track A. Azimov, Ed.
Expires: January 28, 2021 Yandex
July 27, 2020

Methods for Detection and Mitigation of BGP Route Leaks
draft-ietf-grow-route-leak-detection-mitigation-03

Abstract

Problem definition for route leaks and enumeration of types of route leaks are provided in RFC 7908. This document describes a new well-known Large Community that provides a way for route-leak prevention, detection, and mitigation. The configuration process for this Community can be automated with the methodology for setting BGP roles that is described in ietf-idr-bgp-open-policy draft.

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 28, 2021.

Copyright Notice

Copyright (c) 2020 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 (https://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

RFC 7908 [RFC7908] provides a definition of the route-leak problem and enumerates several types of route leaks. For this document, the definition that is applied is that a route leak occurs when a route received from a transit provider or a lateral peer is forwarded (against commonly used policy) to another transit provider or a lateral peer. The commonly used policy is that a route received from a transit provider or a lateral peer MAY be forwarded only to customers.

This document describes a solution for prevention, detection and mitigation route leaks which is based on conveying route-leak detection information in a transitive well-known BGP Large Community. The configuration process for the Large Community MUST be defined according to peering relations between ISPs. This process can be automated with the methodology for setting BGP roles that is described in [I-D.ietf-idr-bgp-open-policy].

The techniques described in this document can be incrementally deployed. If a pair of ISPs and/or Internet Exchanges (IXes) deploy the proposed techniques, then they would detect and mitigate any route leaks that occur in an AS path between them even when other ASes in the path are not upgraded.

2. Peering Relationships

As described in [I-D.ietf-idr-bgp-open-policy] there are several common peering relations between eBGP neighbors:

- o Provider - sender is a transit provider of the neighbor;
- o Customer - sender is a customer of the neighbor;
- o Route Server (RS) - sender is route server at an internet exchange (IX)
- o RS-client - sender is client of an RS at an IX
- o Peer - sender and neighbor are lateral (non-transit) peers;

If a route is received from a provider, peer or RS-client, it MUST follow the 'down only' rule, i.e., it MAY be advertised only to customers. If a route is sent to a customer, peer or RS-client, it also MUST follow the 'down only' rule at each subsequent AS in the AS path.

A standardized transitive route-leak detection signal is needed that will prevent Autonomous Systems (ASes) from leaking and also inform a remote ISP (or AS) in the AS path that a received route violates 'down only' policy. This signal would facilitate a way to stop the propagation of leaked prefixes.

To improve reliability and cover for non-participating preceding neighbor, the signal should be set on both receiver and sender sides.

3. Community vs Attribute

This section presents a brief discussion of the advantages and disadvantages of communities and BGP path attributes for the purpose of route-leak detection.

A transitive path attribute is a native way to implement the route-leak detection signal. Based on the way BGP protocol works, the use of a transitive attribute makes it more certain that the route-leak detection signal would pass unaltered through non-participating (i.e., not upgraded) BGP routers. The main disadvantage of this approach is that the deployment of a new BGP attribute requires a software upgrade in router OS which may delay wide adoption for years.

On the other hand, BGP Communities do not require a router OS update. The potential disadvantage of using a Community for the route-leak

detection signal is that it is more likely to be dropped somewhere along the way in the AS path. Currently, the use of BGP Communities is somewhat overloaded. BGP Communities are already used for numerous applications: different types of route marking, route policy control, blackholing, etc. It is observed that some ASes seem to purposefully or accidentally remove transitive communities on receipt, sometimes well-known ones. Perhaps this issue may be mitigated with strong policy guidance related to the handling of Communities.

Due to frequently occurring regional and global disruptions in Internet connectivity, it is critical to move forward with a solution that is viable in the near term. That solution would be route-leak detection using BGP Community.

Large Communities have much higher capacity, and therefore they are likely to be less overloaded. Hence, Large Community is proposed to be used for route-leak detection. This document suggests reserving <TBD1> class for the purpose of transitive well-known Large Communities that MUST NOT be stripped on ingress or egress.

While it is not a part of this document, the route-leak detection signal described here can also be carried in a transitive BGP Path Attribute, and similar prevention and mitigation techniques as described here would apply (see [I-D.ietf-idr-bgp-open-policy]).

4. Down Only Community

This section specifies the semantics of route-leak detection

Community and its usage. This Community is given the specific name Down Only (DO) Community. The DO Community is carried in a BGP Large Community with a format as shown in Figure 1.

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   TBD1 (class for transitive well-known Large Communities)    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                   TBD2 (subclass for DO)                      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                             ASN                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 1: Format of the DO Community using a Large Community [RFC8092].

The authors studied different options for route-leak mitigation. The main options considered are (1) drop detected route leaks and (2) deprioritize detected route leaks. It can be demonstrated that the loose mode that uses deprioritization is not safe. Traffic Engineering (TE) techniques which limit prefix visibility are quite common. It may happen that a more specific TE prefix is sent only to downstream ASes or to IX(es)/selected peers, and a control Community is used to restrict its propagation. If such a more specific prefix is leaked, deprioritization will not stop such a route leak from propagating. In addition, propagation of leaked prefixes based on deprioritization may result in priority loops leading to BGP wedgies [RFC4264] or even persistent route oscillations.

So, the only truly safe way to implement route-leak mitigation is to drop detected route leaks. The ingress and egress policies corresponding to 'drop detected route leaks' is described in Section 4.1. This policy SHOULD be used as a default behavior.

Nevertheless, early adopters might want to deploy only the signaling and perhaps use it only for diagnostics before applying any route-leak mitigation policy. They are also encouraged to use slightly limited marking, which is described in Section 4.2.

4.1. Route-Leak Mitigation

This section describes the eBGP ingress and egress policies that MUST be used to perform route-leak prevention, detection and mitigation using the DO Community.

The ingress policy MUST use the following procedure:

1. If a route with DO Community set (i.e., DO is attached) is received from a Customer or RS-client, then it is a route leak and MUST be rejected. The procedure halts.

2. If a route with DO Community set is received from Peer (non-transit) and DO value is not equal to the sending neighbor's ASN, then it is a route leak and MUST be rejected. The procedure halts.

3. If a route is received from a Provider, Peer or RS, then the DO Community MUST be added with a value equal to the sending neighbor's ASN.

The egress policy MUST use the following procedure:

1. A route with DO Community set MUST NOT be sent to Providers, Peers, and RS.

2. If a route is sent to a Customer or Peer, then the DO Community MUST be added with a value equal to the ASN of the sender.

The above procedures comprehensively provide route-leak prevention, detection and mitigation. Policy consisting of these procedures SHOULD be used as a default behavior.

4.2. Only Marking

This section describes eBGP ingress and egress marking policies that MUST be used if an AS is not performing route-leak mitigation (i.e., dropping detected route leaks) as described in Section 4.1, but wants to use only marking with DO Community. The slightly limited DO marking (compared to that in Section 4.1) described below guarantees that this DO marking will not limit the leak detection opportunities for subsequent ASes in the AS path.

The ingress policy MUST use the following procedure:

1. If a route with DO Community set is received from a Customer or RS-client, then it is a route leak. The procedure halts.

2. If a route with DO Community set is received from a Peer and DO value is not equal to the sending neighbor's ASN, then it is a route leak. The procedure halts.

3. If a route is received from a Provider, Peer or RS, then the DO Community MUST be added with value equal to the sending neighbor's ASN.

The egress policy MUST use the following procedure:

1. If a route is sent to a Customer or RS-client, then the DO Community MUST be added with value equal to the ASN of the sender.

2. If DO Community is not set and the route is sent to a Peer, then the DO Community MUST be added with value equal to the ASN of the sender.

These above procedures specify setting DO signal in a way that can be used to evaluate the potential impact of route-leak mitigation policy before deploying strict dropping of detected route leaks.

5. Implementation Considerations

It was observed that the majority of BGP implementations do not support negative match for communities like a:b:!c. Considering that it is suggested to replace the second rule from ingress policy with the following:

If a route with DO Community set is received from a Peer and DO value is equal to the sending neighbor's ASN, then it is a valid route, otherwise it is a route leak. The procedure halts.

This rule is based on a weaker assumption that a peer that is doing marking is also doing filtering (dropping detected leaks). That is why networks that do not follow the route-leak mitigation policy in Section 4.1 MUST carefully follow marking rules described in Section 4.2.

6. IANA Considerations

The draft suggests to reserve a Global Administrator ID <TBD1> for transitive well-known Large Community registry. IANA is requested to register a subclass <TBD2> for DO Community in this registry.

7. Security Considerations

In specific circumstances in a state of partial adoption, route-leak mitigation mechanism can result in Denial of Service (DoS) for the victim prefix. Such a scenario may happen only for a prefix that has a single path from the originator to a Tier-1 ISP and only when the prefix is not covered with a less specific prefix with multiple paths to the Tier-1 ISP. If, in such unreliable topology, route leak is injected somewhere inside this single path, then it may be rejected by upper layer providers in the path, thus limiting prefix visibility. While such anomaly is unlikely to happen, such an issue should be easy to debug, since it directly affects the sequence of originator's providers.

With the use of BGP Community, there is often a concern that the Community propagates beyond its intended perimeter and causes harm [streibelt]. However, that concern does not apply to the DO Community because it is a transitive Community that must propagate as far as the update goes.

8. Informative References

[I-D.ietf-idr-bgp-open-policy]

Azimov, A., Bogomazov, E., Bush, R., Patel, K., and K. Sriram, "Route Leak Prevention using Roles in Update and Open messages", draft-ietf-idr-bgp-open-policy-13 (work in progress), July 2020.

[RFC4264] Griffin, T. and G. Huston, "BGP Wedgies", RFC 4264, DOI 10.17487/RFC4264, November 2005, <https://www.rfc-editor.org/info/rfc4264>.

[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>.

[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>.

[streibelt] Streibelt et al., F., "BGP Communities: Even more Worms in the Routing Can", ACM IMC, October 2018, <https://archive.psg.com//181101.imc-communities.pdf>.

Acknowledgements

The authors wish to thank John Scudder, Susan Hares, Ruediger Volk, Jeffrey Haas, Mat Ford, Greg Skinner for their review and comments.

Contributors

The following people made significant contributions to this document and should be considered co-authors:

Brian Dickson
Independent
Email: brian.peter.dickson@gmail.com

Doug Montgomery
USA National Institute of Standards and Technology
Email: dougm@nist.gov

Keyur Patel
Arrcus
Email: keyur@arrcus.com

Andrei Robachevsky
Internet Society
Email: robachevsky@isoc.org

Eugene Bogomazov
Qrator Labs
Email: eb@qrator.net

Randy Bush
Internet Initiative Japan
Email: randy@psg.com

Authors' Addresses

Kotikalapudi Sriram (editor)
USA National Institute of Standards and Technology
100 Bureau Drive
Gaithersburg, MD 20899
United States of America

Email: ksriram@nist.gov

Alexander Azimov (editor)
Yandex
Ulitsa Lva Tolstogo 16
Moscow 119021
Russia

Email: a.e.azimov@gmail.com