IDR
Internet-Draft
Intended status: Standards Track
Expires: April 4, 2018
October 1, 2017

J. Heitz
Cisco
J. Snijders
NTT

BGP Route Leak Protection Community
draft-heitz-idr-route-leak-community-00

Abstract

In general, BGP autonomous system (AS) relationships are either customer-transit or peer-peer. If an AS sends a route received from a transit or a peer to another transit or to another peer, it is considered a route leak. AS relationships are sometimes different for different routes or in different regions. A method of detecting route leaks is proposed that does not require participation by the leaking AS or by IXPs. Only the ASes that perform leak detection need to adopt the proposal. ASes that request leak protection need to send a community to make the request. The proposal works even if the leaking AS or other ASes modify or discard path attributes in the route or create more specific routes.

Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119].

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

In general, BGP autonomous system (AS) relationships are either customer-transit or peer-peer. A route received from a transit or a peer can only be sent to a customer. If an AS sends such a route to a transit or to a peer, then it is considered a route leak. An AS may act as transit for some routes, but not others or in some regions, but not in others. Thus, AS relationships are sometimes different for different routes or in different regions.

An IXP does not add its ASN to the AS_PATH when it announces a route. It is not required to declare an AS relationship. Only the clients of an IXP have relationships with each other. If an IXP were to declare a relationship with its clients, then certain client to client relationships would not be possible without being classified as route leaks. Take an example of 3 ASes that are all connecting to each other through a route server at an IXP. AS1 is transit provider for AS2. AS2 is provider for AS3. There is no relationship that the route server can have with AS2 to make all the client relationships possible.

BGP route leaks and hijacks are described in detail in [RFC7908]. That RFC has references to several leaks and hijacks that have occurred. This document proposes solutions to the leaks type 1, 2, 3, 4 and 6. Type 5 is a hijack, which is addressed by RPKI.

A method of detecting route leaks is proposed that does not require participation by the leaking AS or by IXPs. A leaking AS is not required to recognize, set or transfer any new BGP attributes or communities. Only the ASes that request leak protection and ASes that perform leak protection need to adopt the proposal.

The proposed function runs on the BGP speaker that receives the routes. Thus, any leak can be detected and prevented before the leaking route is even installed in the routing table.

2. Concept

This document automates the concept of Peer Locking described in [Peer-Lock] on a per route basis.

When an AS sends a route to a neighbor, it attaches a set of communities to inform the neighbor which ASes it has nominated to be transit providers for that route. It is saying: "If you receive this route from another AS that is your peer or your customer and my ASN is in the AS_PATH, then my ASN can only be preceded by the ASN of one of my nominated transit providers. If you receive the route with any other ASN preceding my ASN, then it is a route leak." When ASN1 precedes ASN2, then the route was sent from AS2 to AS1 and a packet being forwarded along that route is being forwarded from AS1 to AS2. These communities are called Route Leak Protection Communities or RLP.

A receiving AS may pass these RLPs on to a further AS as it passes the route on. For example, AS1 constructs a set of communities to indicate its nominated transit providers. Suppose these are AS2 and AS10 and it passes the route to AS2. Now, AS2 can pass the route with the communities onto AS3. Then AS3 will learn that AS1 has

nominated AS2 and AS10 as its transit providers for that route. AS2 may add its own transit provider nominations to the route as well. When this set of communities is passed on to a third AS like this, then the third AS must trust the second AS. In the example, AS3 must trust AS2. One way to ensure that trust is for the set of communities to be included in the BGPSEC signature [I-D.ietf-sidr-bgpsec-protocol]. How to do this is for further study.

3. More Specific Routes

More specific routes are often sent to specifically targeted neighbors for traffic engineering purposes within those neighbor ASes only. These are particularly serious when they leak, because they will be preferred over competing routes with shorter netmasks. Even if the route with the shorter netmask has a shorter AS_PATH, the longer netmask wins. More specific routes are valid in some ASes. Therefore a valid ROA must exist for such a route. However, in other ASes, the more specific route is invalid. There is no way for RPKI to invalidate this route in the other ASes.

To indicate that the nominated transit providers are applicable to all routes with a longer netmask than the named route and covered by it, a different community value is used. Such a community is called a Covering RLP or CRLP. It is possible to attach an RLP to a route and attach a different CRLP to the same route. This allows one region of validity to be specified for a route and a different region of validity to be specified for its more specifics.

4. Terminology

Regular Community - BGP Community as defined in [RFC1997].

Large Community - BGP Large Community as defined in [RFC8092].

de-aggregate - A de-aggregate of a first route is a route that has a longer netmask than the first route and is covered by the first route. For example 11::/16 and 12::/16 are de-aggregates of 10::/12, but 1::/16 is not. This is also called a more specific route.

RLP - Route Leak Protection Community. This may be encoded in a regular Community or a Large Community.

RLP Set - All the RLPs attached to one route with the same Nominating ASN. It indicates all the transit ASes

that the Nominating AS has nominated for the given route. If the same route is received from another BGP speaker (also called a path) then the RLPs attached to it do not belong to the same set as those of the first route.

CRLP - Covering RLP. An RLP that applies to the routes that are de-aggregates of the route to which it is attached.

AS - BGP Autonomous System.

ASN - AS Number.

AS_PATH - The AS_PATH as defined in [RFC4271], [RFC6793] and [RFC5065]. Before the AS_PATH is used in this document, confed segments and as-sets are removed and duplicate ASNs are removed.

Neighbor ASN - The last ASN in the AS_PATH of the route. This is usually the ASN of the EBGP speaker from which the route was received. If the route was received from an IXP, then the ASN of the sending BGP speaker is different.

IXP - Internet Exchange Provider. For the purposes of this document, this is an AS that does not add its ASN to the AS_PATH of routes that it announces.

RPKI - A method of IP prefix origin AS validation. Described in [RFC6811] and other RFCs.

ROA - Route Origin Authorization. A signed record linking IP addresses to an AS. Used by RPKI. Described in [RFC6482]

5. Encoding

A nomination of transit provider is encoded in a BGP Large Community as follows:

```
     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
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |                            RLP Code                           |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |                         Nominating ASN                        |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |                          Nominated ASN                        |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

The fields are as shown below:

RLP Code - A 32 bit Autonomous System Number to indicate that this is a Route Leak Protection Large Community. Either one of two values is used. The first indicates that the community applies to the attached route. The second value indicates that the community applies to all routes with a longer netmask that are covered by the attached route. The first value indicates an RLP and the second indicates a CRLP. Both values are to be assigned by IANA from the BGP ASN registry.

Nominating ASN - ASN of the AS that is nominating a transit ASN.

Nominated ASN - ASN of the transit ASN being nominated.

An AS MUST attach an RLP Large community for every ASN that it is nominating as a transit ASN. To indicate that it is nominating no transit ASNs, an AS attaches a single RLP Large Community with a Nominated ASN of 0. An AS that is not declaring its transit ASNs does not attach any RLP Large Communities with its own ASN as Nominating ASN.

5.1. Limited Alternative using Regular Communities

As an alternative to BGP Large Communities, regular BGP communities can be used. However, this will only work to nominate 2-octet transit ASNs and it cannot be passed onto subsequent ASes. The values to use in the regular community are as follows:

```
     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
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |            RLP Code           |        Nominated ASN          |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

The fields are as shown below:

RLP Code- A 16 bit Number to indicate that this is a Route Leak Protection Community. Either one of two values is used. The first indicates that the community applies to the attached route. The second value indicates that the community applies to all routes with a longer netmask that are covered by the attached route. The first value indicates an RLP and the second indicates a CRLP. The values are to be agreed upon by the neighboring ASes.

Nominated ASN - ASN of the transit ASN being nominated.

6. Procedures

If an RLP or a CRLP is received in the form of a regular community, then it is converted into an equivalent Large Community before being used. The Nominating ASN is set to the Neighbor ASN in the AS_PATH. Using the neighbor ASN in the AS_PATH rather than the ASN of the neighbor router allows the community to pass through an IXP route server.

If the Nominating ASN in an RLP or CRLP does not appear in the AS_PATH of the route to which it is attached, then the RLP or CRLP is discarded.

Whenever the RLPs or CRLPs applicable to a route change and that route was received from either a peer AS or a customer AS, the following procedure is executed.

A BGP speaker may have received several routes to the same prefix from multiple neighbors. All of the RLPs that have been received in all those routes are collected together. The RLPs are collected from all the received routes for the prefix, not just the bestpath. The RLPs from a just received route are also collected unless they are explicitly denied by policy. An AS may locally create an RLP set and collect it too. Covering RLPs are also collected from covering routes. The RLPs for a prefix are grouped by neighbor ASN and nominating ASN.

A subset of this collection of RLPs is used to validate the route. The subset to use is determined as follows:

If the operator has created a local set of RPLs, then that set is used. The operator may add RLPs received from other sources as per local policy.

Else if an RLP set exists that has the Nominating ASN equal to the Neighbor ASN, then only this RLP set is used.

Else if a CRLP set exists that has the Nominating ASN equal to the Neighbor ASN, then only this CRLP set is used. If there are multiple such CRLP sets with different netmask lengths, then the set with the longest netmask length is used.

Else if at least one RLP set exists, then the union of all RLP sets is used.

Else if CRPL sets exist, then the union of the sets with the longest netmask in the associated route is used.

Note that if the used RLP sets differ, then some of them cannot be trusted and should not have been accepted when the associated route was received.

Next, the Nominating ASN is found in the AS_PATH of the route. If the ASN preceding the Nominating ASN on the AS_PATH is not equal to one of the Nominated ASNs in the RLP set, then the route is a leak. The response to a leak is a local decision. Some possible actions are to assign a low LOCAL_PREF to the route or not to install the route in the Loc-RIB or to drop the route.

7. Deployment Considerations

If an AS attaches an RLP set to a route with its own ASN as the Nominating ASN and it announces that route to multiple BGP speakers, then it MUST either attach the same RLP set or no RLP set to the announcements sent to each speaker. It MUST NOT attach a different RLP set to the same route announced to different BGP speakers.

An AS MUST NOT remove any RLPs from an RLP set that it has received when forwarding the RLP set to another AS, except if the Nominated ASN is 0. However, it MAY delete the complete RLP set. An AS MAY add an RLP to an RLP set with its own ASN as the Nominated ASN.

The same considerations apply to CRLPs.

A route wih an attached RLP may be discarded because it is withdrawn or because it is invalidated by another RLP. If that RLP caused a second route to be invalidated and discarded, then a BGP REFRESH message may be issued to recover the second route. If the RLP on a route invalidates the route itself or if a set of routes invalidate each other, then REFRESH messages MUST NOT be issued to recover those routes. A subsequent change in routing policy may independently cause a REFRESH message to be issued.

8. Security Considerations

An AS can attach RLPs with Nominating ASN different to its own ASN in order to falsely cause the routes from another AS to be detected as a leak. For this reason, RLPs should only be accepted from trusted ASes. If the Nominating ASN in an RLP is equal to the Neighbor ASN and the Neighbor ASN can be verified, then the RLP can be trusted. In other words, if an AS declares incorrect transits for itself, then it is hurting only itself.

RLPs disclose which ASes are the Nominating AS's transit providers. This may be sensitive information for some. However, for another AS to detect a route leak, it needs to know this information. This concern can be mitigated by sending RLPs to transit providers only, not to peers and customers. This is just telling one's transit provider not to block one's route from one's other transit providers. In that case it is not a concern. If an AS does not want to disclose its transits, then it is only not requesting route leak protection, it is not affecting route leak protection for any other AS.

9. IANA Considerations

IANA is requested to assign an ASN for the RLP identifier and the CRLP identifier from the BGP ASN registry.

10. Acknowledgments

Juan Alcaide, Kalpesh Zinjuwadia.

11. Discussion Topics

11.1. Use of the Regular Community

This is of limited use and only until Large Communities are widespread. Since the use of Regular Communities for RLP is by private agreement between neighboring ASes only, there is no need to standardize it.

11.2. Limited Reach

The RLP is really only useful in the first 2 or 3 AS hops. After it has traveled 10 ASes, it is only using space. One way to determine how many AS hops an RLP has traveled is to find the Nominating ASN in the AS_PATH.

11.3. Well Known Large Communities

The value in the first 4 octets of the Large Community that indicates an RLP or CRLP is taken from the ASN registry. An alternative is to define a range of ASNs to be used for future well known Large Communities.

12. References

12.1. Normative 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>.

[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>.

[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>.

[RFC6793] Vohra, Q. and E. Chen, "BGP Support for Four-Octet Autonomous System (AS) Number Space", RFC 6793, DOI 10.17487/RFC6793, December 2012, <https://www.rfc-editor.org/info/rfc6793>.

[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>.

12.2. Informative References

[I-D.ietf-sidr-bgpsec-protocol] Lepinski, M. and K. Sriram, "BGPsec Protocol Specification", draft-ietf-sidr-bgpsec-protocol-23 (work in progress), April 2017.

[Peer-Lock]
Snijders, J., "Peer Locking", June 2016, <https://www.nanog.org/sites/default/files/Snijders_Everyday_Practical_Bgp.pdf>.

[RFC6482] Lepinski, M., Kent, S., and D. Kong, "A Profile for Route Origin Authorizations (ROAs)", RFC 6482, DOI 10.17487/RFC6482, February 2012, <https://www.rfc-editor.org/info/rfc6482>.

[RFC6811] Mohapatra, P., Scudder, J., Ward, D., Bush, R., and R. Austein, "BGP Prefix Origin Validation", RFC 6811, DOI 10.17487/RFC6811, January 2013, <https://www.rfc-editor.org/info/rfc6811>.

[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>.

Authors' Addresses

Jakob Heitz
Cisco
170 West Tasman Drive
San Jose, CA 95134
USA

Email: jheitz@cisco.com

Job Snijders
NTT Communications
Theodorus Majofskistraat 100
Amsterdam 1065 SZ
The Netherlands

Email: job@ntt.net