IDR and SIDR
Internet-Draft
Intended status: Standards Track
Expires: September 6, 2018

K. Sriram
D. Montgomery
USA NIST
B. Dickson

K. Patel
Arrcus
A. Robachevsky
Internet Society

March 5, 2018

Methods for Detection and Mitigation of BGP Route Leaks
draft-ietf-idr-route-leak-detection-mitigation-08

Abstract

Problem definition for route leaks and enumeration of types of route leaks are provided in RFC 7908. This document specifies BGP enhancements that significantly extend its route-leak detection and mitigation capabilities. The solution involves carrying a per-hop route-leak protection (RLP) field in BGP updates. The RLP fields are carried in a new optional transitive attribute, called BGP RLP attribute. The RLP attribute helps with detection and mitigation of route leaks at ASes downstream from the leaking AS (in the path of BGP update). This is an inter-AS (multi-hop) solution mechanism. This solution complements the intra-AS (local AS) route-leak avoidance solution that is described in ietf-idr-bgp-open-policy draft.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

RFC 7908 [RFC7908] provides a definition of the route leak problem, and also enumerates several types of route leaks. This document first examines which of those route-leak types are detected and mitigated by the existing Origin Validation (OV) [RFC6811] method. OV [RFC6811] and BGPsec path validation [RFC8205] together offer mechanisms to protect against re-originations and hijacks of IP prefixes as well as man-in-the-middle (MITM) AS path modifications. Route leaks [RFC7908] are another type of vulnerability in the global BGP routing system against which OV offers very limited protection. BGPsec path validation provides cryptographic protection for some aspects of BGP update messages, but in its current form BGPsec does not offer any protection against route leaks.

For the types of route leaks enumerated in RFC 7908 [RFC7908], where the OV method does not offer a solution, this document specifies BGP enhancements that significantly extend its route-leak detection and mitigation capabilities. The solution involves carrying a per-hop route-leak protection (RLP) field in BGP updates. The RLP fields are carried in a new optional transitive attribute, called BGP RLP attribute. The RLP attribute helps with detection and mitigation of route leaks at ASes downstream from the leaking AS (in the path of BGP update). This is an inter-AS (multi-hop) solution mechanism. This solution complements the intra-AS (local AS) route-leak avoidance solution that is described in [I-D.ietf-idr-bgp-open-policy].

The RLP mechanism is backward compatible with BGP routers that are not upgraded to perform RLP. Early adopters would see significant benefits. If a group of big ISPs deploy RLP, then they would be helping each other by blocking route leaks originated within one's customer cone from propagating into a peer's AS or their customer cone. The intra-AS (local AS) route-leak avoidance solution [I-D.ietf-idr-bgp-open-policy] works to prevent a local AS from leaking routes but requires 100% deployment in order to prevent propagation of route leaks across AS boundaries. Hence, the inter-AS RLP solution (this document) and the intra-AS solution [I-D.ietf-idr-bgp-open-policy] are complementary.

The inter-AS RLP solution is meant to be initially implemented as an enhancement of BGP without requiring BGPsec. However, when BGPsec is deployed in the future, the solution should be incorporated in BGPsec, enabling cryptographic protection for the RLP fields. That is one way of securing the solution against malicious route leaks.

It is not claimed that the RLP solution detects all possible types of route leaks [RFC7908], but it detects several types (see Section 2), especially considering some significant route-leak occurrences that have been observed in recent years. The RLP solution mechanism is described in Section 3. A review of related prior work is presented in Appendix A. An intra-AS route-leak prevention method using BGP Community is discussed in Appendix B. The document also includes a stopgap method for detection and mitigation of route leaks for an intermediate phase when OV is deployed but BGP protocol on the wire is unchanged (see Appendix C). Design rationale and discussion are presented in Appendix D.

2. Rote-Leak Types that the Solution Must Address

Referring to the enumeration of route leaks discussed in [RFC7908], Table 1 summarizes the route-leak detection capability offered by OV and BGPsec for different types of route leaks.

A detailed explanation of the contents of Table 1 is as follows. It is readily observed that route leaks of Types 1, 2, 3, and 4 are not detected by OV or BGPsec in its current form. Clearly, Type 5 route leak involves re-origination or hijacking, and hence can be detected by OV. In the case of Type 5 route leak, there would be no existing ROAs to validate a re-originated prefix or more specific, but instead a covering ROA would normally exist with the legitimate AS, and hence the update will be considered Invalid by OV.

Type of Route Leak	Current State of Detection Coverage
Type 1: Hairpin Turn with Full Prefix	Neither OV nor BGPsec (in its current form) detects Type 1.
Type 2: Lateral ISP-ISP-ISP Leak	Neither OV nor BGPsec (in its current form) detects Type 2.
Type 3: Leak of Transit-Provider Prefixes to Peer	Neither OV nor BGPsec (in its current form) detects Type 3.
Type 4: Leak of Peer Prefixes to Transit Provider	Neither OV nor BGPsec (in its current form) detects Type 4.
Type 5: Prefix Re-Origination with Data Path to Legitimate Origin	OV detects Type 5.
Type 6: Accidental Leak of Internal Prefixes and More Specifics	For internal prefixes never meant to be routed on the Internet, OV helps detect their leak; they might either have no covering ROA or have an AS0-ROA to always filter them. In the case of accidental leak of more specifics, OV may offer some detection due to ROA maxLength.

Table 1: Examination of Route-Leak Detection Capability of Origin Validation and Current BGPsec Path Validation

In the case of Type 6 leaks involving internal prefixes that are not meant to be routed in the Internet, they are likely to be detected by OV. That is because such prefixes might either have no covering ROA or have an AS0-ROA to always filter them. In the case of Type 6 leaks that are due to accidental leak of more specifics, they may be detected due to violation of ROA maxLength. BGPsec (i.e., path validation) in its current form does not detect Type 6. However, route leaks of Type 6 are least problematic due to the following reasons. In the case of leak of more specifics, the offending AS is itself the legitimate destination of the leaked more-specific prefixes. Hence, in most cases of this type, the data traffic is not misrouted. Also, leaked announcements of Type 6 are short-lived and typically withdrawn quickly following the announcements. Further, the MaxPrefix limit may kick-in in some receiving routers and that

helps limit the propagation of sometimes large number of leaked routes of Type 6.

Realistically, BGPsec may take a much longer time being deployed than OV. Hence, solution proposals for route leaks should consider both scenarios: (A) OV only (without BGPsec) and (B) OV plus BGPsec. Assuming an initial scenario A, and based on the above discussion and Table 1, it is evident that the solution method should focus primarily on route leaks of Types 1, 2, 3, and 4.

3. Mechanisms for Detection and Mitigation of Route Leaks

There are two considerations for route leaks: (1) Prevention of route leaks from a local AS [I-D.ietf-idr-bgp-open-policy], and (2) Detection and mitigation of route leaks in ASes that are downstream from the leaking AS (in the path of BGP update). This document focuses on the latter, and the details of the mechanism are described in this section.

3.1. Ascertaining Peering Relationship

There are four possible peering relationships (i.e., roles) an AS can have with a neighbor AS: (1) Provider: transit-provider for all prefixes exchanged, (2) Customer: customer for all prefixes exchanged, (3) Lateral Peer: lateral peer (i.e., non-transit) for all prefixes exchanged, and (4) Complex: different relationships for different sets of prefixes [Luckie]. On a per-prefix basis, the peering role types simplify to provider, customer, or lateral peer.

Operators rely on some form of out-of-band (OOB) (i.e., external to BGP) communication to exchange information about their peering relationship, AS number, interface IP address, etc. If the relationship is complex, the OOB communication also includes the sets of prefixes for which they have different roles. [I-D.ietf-idr-bgp-open-policy] introduces a method of re-confirming the BGP Role during BGP OPEN messaging (except when the role is complex). It defines a new BGP Role capability, which helps in re-confirming the relationship when it is provider, customer, or lateral peer. BGP Role does not replace the OOB communication since it relies on the OOB communication to set the role type in the BGP OPEN message. However, BGP Role provides a means to double check, and if there is a contradiction detected via the BGP Role messages, then a Role Mismatch Notification is sent [I-D.ietf-idr-bgp-open-policy].

When the BGP relationship information has been correctly exchanged (i.e., free of contradictions) including the sets of prefixes with different roles (if complex), then this information SHOULD be used to set the role per-prefix with each peer. For example, if the local

AS's role is Provider with a neighbor AS, then the per-prefix role is set to 'Provider' for all prefixes sent to the neighbor, and set to 'Customer' for all prefixes received from the neighbor.

Once the per-prefix roles are set, this information is used in the RLP solution mechanism that is described in Section 3.2 and Section 3.3.

3.2. Route-Leak Protection (RLP) Field Encoding by Sending Router

The key principle is that, in the event of a route leak, a receiving router in a transit-provider AS (e.g., referring to Figure 1, ISP2 (AS2) router) should be able to detect from the update message that its customer AS (e.g., AS3 in Figure 1) should not have forwarded the update (towards the transit-provider AS). This means that at least one of the ASes in the AS path of the update signaled that it sent the update to its customer or lateral peer, but forbade any subsequent 'Up' (customer to provider) or 'Lateral' (peer to peer) forwarding. This signaling is achieved by a Route-Leak Protection (RLP) field as described below.

```
                                      /\              /\
                                       \ route-leak(P)/
                                        \ propagated /
                                         \          /
              +------------+    peer    +------------+
        ______| ISP1 (AS1) |----------->|  ISP2 (AS2)|---------->
       /       ------------+  prefix(P) +------------+ route-leak(P)
      | prefix |          \   update      /\        \  propagated
       \  (P)  /           \              /          \
        -------   prefix(P) \            /            \
                     update  \          /              \
                              \        /route-leak(P)  \/
                              \/      /
                           +---------------+
                           | customer(AS3) |
                           +---------------+
```

Figure 1: Illustration of the basic notion of a route leak.

For the purpose of route-leak detection and mitigation, the RLP field value MUST be set to one of two values as follows:

o 0: This is the default value (i.e., "nothing specified"),

- 1: This is the 'Do not Propagate Up or Lateral' indication; sender indicating that the route must not be forwarded 'Up' towards a transit-provider AS or to a lateral (i.e., non-transit) peer AS.

The RLP indications are set on a per prefix basis. This is because some peering relations between neighbors can be complex (see Section 3.1). Further, the RLP indications are also set on a per hop (i.e., per AS) basis.

There are two different scenarios when a sending AS MUST set value 1 in the RLP field: (a) when sending the update to a customer AS, and (b) when sending the update to a lateral peer (i.e., non-transit) AS. In essence, in both scenarios, the intent of RLP = 1 is that the neighbor AS and any receiving AS along the subsequent AS path SHOULD NOT forward the update 'Up' towards its (receiving AS's) transit-provider AS or laterally towards its peer AS.

When sending an update 'Up' to a transit-provider AS, the RLP encoding MUST be set to the default value of 0. When a sending AS sets the RLP encoding to 0, it is indicating to the receiving AS that the update can be propagated in any direction (i.e., towards transit-provider, customer, or lateral peer).

The two-state specification in the RLP field (as described above) works for detection and mitigation of route leaks of Types 1, 2, 3, and 4 which are the focus here (see Section 3.3 and Section 3.4).

An AS MUST NOT rewrite/reset the values set by any preceding ASes in their respective RLP fields.

The RLP encoding MUST be carried in BGP-4 [RFC4271] updates in a new BGP optional transitive attribute (see Section 3.2.1). In BGPsec, it must be carried in the Flags field (see Section 3.2.2).

In partial deployment, there may be eBGP routers in the AS path that are not upgraded and hence do not participate in RLP. However, the RLP mechanism is backward compatible. Participating ASes can detect and mitigate route leaks while ASes not upgraded to do RLP would likely allow route leaks to propagate. If big ISPs deploy RLP, then they would be helping each other by not allowing route leaks originated within one's customer cone to propagate into another's AS or their customer cone. This accords significant benefit to early adopters.

3.2.1. BGP RLP Attribute

The BGP RLP attribute is a new BGP optional transitive attribute. The attribute type code for the RLP attribute is to be assigned by IANA. The length field of this attribute is 2 octets. The value field of the RLP attribute is defined as a set of one or more pairs of ASN (4 octets) and RLP (one octet) fields as described below (Figure 2).

```
+-----------------------+ -\
| ASN: N                |   |
+-----------------------+    >  (Most recently added)
| RLP: N                |   |
+-----------------------+ -/
 ...........
+-----------------------+ -\
| ASN: 1                |   |
+-----------------------+    >  (Least recently added)
| RLP: 1                |   |
+-----------------------+ -/
```

Figure 2: BGP RLP Attribute format.

The RLP Attribute value is a sequence of these two components (see Figure 2):

ASN: Four octets encoding the public registered AS number of a BGP speaker.

RLP Field: One octet encoding the RLP Field bits. The value of the RLP Field octet can be 0 (decimal) or 1 (decimal) as described above in Section 3.2.1. Its usage will be further discussed in subsequent sections.

If all ASes in the AS_PATH of a route are upgraded to participate in RLP, then the ASNs in the RLP TLV in Figure 2 will correspond one-to-one with sequence of ASes in the AS_PATH (excluding prepends). If some ASes do not participate, then one or more {ASN, RLP} tuples may be missing in the RLP attribute relative to the AS_PATH.

3.2.2. Carrying RLP Field Values in the BGPsec Flags

In BGPsec-enabled routers that are also performing RLP, the RLP encoding MUST be accommodated in the existing Flags field in BGPsec updates. The Flags field is part of the Secure_Path Segment in BGPsec updates [RFC8205]. It is one octet long, and one Flags field

is available for each AS hop, and currently only the first bit is used in BGPsec. So, there are 7 bits that are currently unused in the Flags field. One of these bits can be designated for the RLP field value (see Section 3.2.1). This bit can be set to 0 when the RLP Field value is 0 and set to 1 when the RLP Field value is 1. Since the BGPsec protocol specification requires a sending AS to include the Flags field in the data that are signed over, the RLP field for each hop (assuming it would be part of the Flags field as described) will be protected under the sending AS's signature.

3.3. Recommended Actions at a Receiving Router for Detection of Route Leaks

The following receiver algorithm is RECOMMENDED for detecting route leaks:

A receiving router MUST mark an update as a 'Route Leak' if ALL of the following conditions hold true:

1. The update is received from a customer or lateral peer AS.

2. The update has the RLP Field set to 1 (i.e., 'Do not Propagate Up or Lateral') indication for one or more hops (excluding the most recent) in the AS path.

The reason for stating "excluding the most recent" in the above algorithm is as follows. An ISP should look at RLP values set by ASes preceding the immediate sending AS in order to ascertain a leak. The receiving router already knows that the most recent hop in the update is from its customer or lateral-peer AS to itself, and it does not need to rely on the RLP field value set by that AS (i.e., the immediate neighbor AS in the AS path) for detection of route leaks.

If the RLP encoding is secured by BGPsec (see Section 3.2) and hence protected against tampering by intermediate ASes, then there would be added certainty in the route-leak detection algorithm described above (see discussions in Appendix D.1 and Appendix D.2).

3.4. Possible Actions at a Receiving Router for Mitigation

After applying the above detection algorithm, a receiving router may use any policy-based algorithm of its own choosing to mitigate any detected route leaks. An example receiver algorithm for mitigating a route leak is as follows:

- If an update from a customer or lateral peer AS is marked as a 'Route Leak' (see Section 3.3), then the receiving router SHOULD prefer an alternate unmarked route.

- o If no alternate unmarked route is available, then a route marked as a 'Route Leak' MAY be accepted.

A basic principle here is that if an AS receives and marks a customer route as 'Route Leak', then the AS should override the "prefer customer route" policy, and instead prefer an alternate 'clean' route learned from another customer, a lateral peer, or a transit provider. This can be implemented by adjusting the local preference for the routes in consideration.

4. Security Considerations

The Route-Leak Protection (RLP) field requires cryptographic protection in order to prevent malicious route leaks. In the future, in conjunction with BGPsec deployment, the RLP field will be included in the Flags field in the Secure_Path Segment in BGPsec updates. So, the cryptographic security mechanisms in BGPsec will also apply to the RLP field. The reader is therefore directed to the security considerations provided in [RFC8205].

5. IANA Considerations

IANA is requested to register a new optional, transitive BGP Path Attribute, named "Route Leak Protection" in the BGP Path Attributes registry. The attribute type code is TBD. The reference for this new attribute is this document (i.e., the RFC that replaces this draft). The length field of this attribute is 2 octets, and the length of the value field of this attribute is variable (see Figure 2) in Section 3.2.1 of this document).

6. References

6.1. Normative References

[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>.

6.2. Informative References

[Anwar] Anwar, R., Niaz, H., Choffnes, D., Cunha, I., Gill, P., and N. Katz-Bassett, "Investigating Interdomain Routing Policies in the Wild", ACM Internet Measurement Conference (IMC), October 2015, <http://www.cs.usc.edu/assets/007/94928.pdf>.

[Cowie2010]
Cowie, J., "China's 18 Minute Mystery", Dyn Research/Renesys Blog, November 2010, <http://research.dyn.com/2010/11/chinas-18-minute-mystery/>.

[Cowie2013]
Cowie, J., "The New Threat: Targeted Internet Traffic Misdirection", Dyn Research/Renesys Blog, November 2013, <http://research.dyn.com/2013/11/mitm-internet-hijacking/>.

[draft-dickson-sidr-route-leak-solns]
Dickson, B., "Route Leaks -- Proposed Solutions", IETF Internet Draft (expired), March 2012, <https://tools.ietf.org/html/draft-dickson-sidr-route-leak-solns-01>.

[draft-kunzinger-idrp-ISO10747-01]
Kunzinger, C., "Inter-Domain Routing Protocol (IDRP)", IETF Internet Draft (expired), November 1994, <https://tools.ietf.org/pdf/draft-kunzinger-idrp-ISO10747-01.pdf>.

[Gao] Gao, L. and J. Rexford, "Stable Internet routing without global coordination", IEEE/ACM Transactions on Networking, December 2001, <http://www.cs.princeton.edu/~jrex/papers/sigmetrics00.long.pdf>.

[Gill] Gill, P., Schapira, M., and S. Goldberg, "A Survey of Interdomain Routing Policies", ACM SIGCOMM Computer Communication Review, January 2014, <https://www.cs.bu.edu/~goldbe/papers/survey.pdf>.

[Giotsas] Giotsas, V. and S. Zhou, "Valley-free violation in Internet routing - Analysis based on BGP Community data", IEEE ICC 2012, June 2012.

[Hiran] Hiran, R., Carlsson, N., and P. Gill, "Characterizing Large-scale Routing Anomalies: A Case Study of the China Telecom Incident", PAM 2013, March 2013, <http://www3.cs.stonybrook.edu/~phillipa/papers/CTelecom.html>.

[Huston2012]
Huston, G., "Leaking Routes", March 2012, <http://labs.apnic.net/blabs/?p=139/>.

[Huston2014]
Huston, G., "What's so special about 512?", September 2014, <http://labs.apnic.net/blabs/?p=520/>.

[I-D.ietf-idr-aspath-orf]
Hares, S. and K. Patel, "AS Path Based Outbound Route Filter for BGP-4", draft-ietf-idr-aspath-orf-13 (work in progress), December 2016.

[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-02 (work in progress), January 2018.

[Kapela-Pilosov]
Pilosov, A. and T. Kapela, "Stealing the Internet: An Internet-Scale Man in the Middle Attack", DEFCON-16 Las Vegas, NV, USA, August 2008, <https://www.defcon.org/images/defcon-16/dc16-presentations/defcon-16-pilosov-kapela.pdf>.

[Kephart] Kephart, N., "Route Leak Causes Amazon and AWS Outage", ThousandEyes Blog, June 2015, <https://blog.thousandeyes.com/route-leak-causes-amazon-and-aws-outage>.

[Khare] Khare, V., Ju, Q., and B. Zhang, "Concurrent Prefix Hijacks: Occurrence and Impacts", IMC 2012, Boston, MA, November 2012, <http://www.cs.arizona.edu/~bzhang/paper/12-imc-hijack.pdf>.

[Labovitz]
Labovitz, C., "Additional Discussion of the April China BGP Hijack Incident", Arbor Networks IT Security Blog, November 2010, <http://www.arbornetworks.com/asert/2010/11/additional-discussion-of-the-april-china-bgp-hijack-incident/>.

[LRL] Khare, V., Ju, Q., and B. Zhang, "Large Route Leaks", Project web page, 2012, <http://nrl.cs.arizona.edu/projects/lsrl-events-from-2003-to-2009/>.

[Luckie] Luckie, M., Huffaker, B., Dhamdhere, A., Giotsas, V., and kc. claffy, "AS Relationships, Customer Cones, and Validation", IMC 2013, October 2013, <http://www.caida.org/~amogh/papers/asrank-IMC13.pdf>.

[Madory]   Madory, D., "Why Far-Flung Parts of the Internet Broke Today", Dyn Research/Renesys Blog, September 2014, <http://research.dyn.com/2014/09/why-the-internet-broke-today/>.

[Mauch]   Mauch, J., "BGP Routing Leak Detection System", Project web page, 2014, <http://puck.nether.net/bgp/leakinfo.cgi/>.

[Mauch-nanog]
Mauch, J., "Detecting Routing Leaks by Counting", NANOG-41 Albuquerque, NM, USA, October 2007, <https://www.nanog.org/meetings/nanog41/presentations/mauch-lightning.pdf>.

[Nanog-thread-June2016]
"Intra-AS messaging for route leak prevention", NANOG Email List - Discussion Thread , June 2016, <http://mailman.nanog.org/pipermail/nanog/2016-June/thread.html#86348>.

[NIST-800-54]
Kuhn, D., Sriram, K., and D. Montgomery, "Border Gateway Protocol Security", NIST Special Publication 800-54, July 2007, <http://csrc.nist.gov/publications/nistpubs/800-54/SP800-54.pdf>.

[Paseka]   Paseka, T., "Why Google Went Offline Today and a Bit about How the Internet Works", CloudFare Blog, November 2012, <http://blog.cloudflare.com/why-google-went-offline-today-and-a-bit-about/>.

[proceedings-sixth-ietf]
Gross, P., "Proceedings of the April 22-24, 1987 Internet Engineering Task Force", April 1987, <https://www.ietf.org/proceedings/06.pdf>.

[RFC1105-obsolete]
Lougheed, K. and Y. Rekhter, "A Border Gateway Protocol (BGP)", IETF RFC (obsolete), June 1989, <https://tools.ietf.org/html/rfc1105>.

[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>.

[RFC7454] Durand, J., Pepelnjak, I., and G. Doering, "BGP Operations and Security", BCP 194, RFC 7454, DOI 10.17487/RFC7454, February 2015, <https://www.rfc-editor.org/info/rfc7454>.

[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>.

[RFC8205] Lepinski, M., Ed. and K. Sriram, Ed., "BGPsec Protocol Specification", RFC 8205, DOI 10.17487/RFC8205, September 2017, <https://www.rfc-editor.org/info/rfc8205>.

[Snijders]
Snijders, J., "Practical everyday BGP filtering with AS_PATH filters: Peer Locking", NANOG-47 Chicago, IL, USA, June 2016, <https://www.nanog.org/sites/default/files/Snijders_Everyday_Practical_Bgp.pdf>.

[Sriram]   Sriram, K., Montgomery, D., Dickson, B., Patel, K., and A. Robachevsky , "Methods for Detection and Mitigation of BGP Route Leaks", IETF-95 IDR WG Meeting), April 2016, <https://www.ietf.org/proceedings/95/slides/slides-95-idr-13.pdf>.

[Toonk]   Toonk, A., "What Caused Today's Internet Hiccup", August 2014, <http://www.bgpmon.net/what-caused-todays-internet-hiccup/>.

[Toonk2015-A]
Toonk, A., "What caused the Google service interruption", March 2015, <http://www.bgpmon.net/what-caused-the-google-service-interruption/>.

[Toonk2015-B]
Toonk, A., "Massive route leak causes Internet slowdown", June 2015, <http://www.bgpmon.net/massive-route-leak-cause-internet-slowdown/>.

[Wijchers]
Wijchers, B. and B. Overeinder, "Quantitative Analysis of BGP Route Leaks", RIPE-69, November 2014, <https://ripe69.ripe.net/presentations/157-RIPE-69-Routing-WG.pdf>.

[Zmijewski]
Zmijewski, E., "Indonesia Hijacks the World", Dyn Research/Renesys Blog, April 2014, <http://research.dyn.com/2014/04/indonesia-hijacks-world/>.

Appendix A. Related Prior Work

The solution described in this document is based on setting an attribute in BGP route announcement to manage the transmission/receipt of the announcement based on the type of neighbor (e.g., customer, transit provider, etc.). Documented prior work related to this basic idea and mechanism dates back to at least the 1980's. Some examples of prior work are: (1) Information flow rules described in [proceedings-sixth-ietf] (see pp. 195-196); (2) Link Type described in [RFC1105-obsolete] (see pp. 4-5); (3) Hierarchical Recording described in [draft-kunzinger-idrp-ISO10747-01] (see Section 6.3.1.12). The problem of route leaks and possible solution mechanisms based on encoding peering-link type information, e.g., P2C (i.e., Transit-Provider to Customer), C2P (i.e., Customer to Transit-Provider), p2p (i.e., peer to peer) etc., in BGPsec updates and protecting the same under BGPsec path signatures have been discussed in IETF SIDR WG at least since 2011. [draft-dickson-sidr-route-leak-solns] attempted to describe these mechanisms in a BGPsec context. The draft expired in 2012. [draft-dickson-sidr-route-leak-solns] defined neighbor relationships on a per link basis, but in the current document the relationship is encoded per prefix, as routes for prefixes with different peering relationships may be sent over the same link. Also [draft-dickson-sidr-route-leak-solns] proposed a second signature block for the link type encoding, separate from the path signature block in BGPsec. By contrast, in the current document when BGPsec-based solution is considered, cryptographic protection is provided for Route-Leak Protection (RLP) encoding using the same signature block as that for path signatures (see Section 3.2.2).

Appendix B. Prevention of Route Leaks at Local AS: Intra-AS Messaging

Note: The intra-AS messaging for route leak prevention can be done using a non-transitive BGP Community or Attribute. The Community-based method is described below. For the BGP Attribute-based method, see [I-D.ietf-idr-bgp-open-policy].

B.1. Non-Transitive BGP Community for Intra-AS Messaging

The following procedure (or similar) for intra-AS messaging (i.e., between ingress and egress routers) for prevention of route leaks is a fairly common practice used by large ISPs. (Note: This information

was gathered from discussions on the NANOG mailing list [Nanog-thread-June2016] as well as through private discussions with operators of large ISP networks.)

Routes are tagged on ingress to an AS with communities for origin, including the type of eBGP peer it was learned from (customer, provider or lateral peer), geographic location, etc. The community attributes are carried across the AS with the routes. These communities are used along with additional logic in route policies to determine which routes are to be announced to which eBGP peers and which are to be dropped. In this process, the ISP's AS also ensures that routes learned from a transit-provider or a lateral peer (i.e., non-transit) at an ingress router are not leaked at an egress router to another transit-provider or lateral peer.

Additionally, in many cases, ISP network operators' outbound policies require explicit matches for expected communities before passing routes. This helps ensure that that if an update has been entered into the RIB-in but has missed its ingress community tagging (due to a missing/misapplied ingress policy), it will not be inadvertently leaked.

The above procedure (or a simplified version of it) is also applicable when an AS consists of a single eBGP router. It is recommended that all AS operators SHOULD implement the procedure described above (or similar that is appropriate for their network) to prevent route leaks that they have direct control over.

Appendix C. Stopgap Solution when Only Origin Validation is Deployed

A stopgap method is described here for detection and mitigation of route leaks for the intermediate phase when OV is deployed but BGP protocol on the wire is unchanged. The stopgap solution can be in the form of construction of a prefix filter list from ROAs. A suggested procedure for constructing such a list comprises of the following steps:

- o ISP makes a list of all the ASes (Cust_AS_List) that are in its customer cone (ISP's own AS is also included in the list). (Some of the ASes in Cust_AS_List may be multi-homed to another ISP and that is OK.)

- o ISP downloads from the RPKI repositories a complete list (Cust_ROA_List) of valid ROAs that contain any of the ASes in Cust_AS_List.

- o ISP creates a list of all the prefixes (Cust_Prfx_List) that are contained in any of the ROAs in Cust_ROA_List.

o Cust_Prfx_List is the allowed list of prefixes that is permitted by the ISP's AS, and will be forwarded by the ISP to upstream ISPs, customers, and peers.

o A route for a prefix that is not in Cust_Prfx_List but announced by one of ISP's customers is 'marked' as a potential route leak. Further, the ISP's router SHOULD prefer an alternate route that is Valid (i.e., valid according to origin validation) and 'clean' (i.e., not marked) over the 'marked' route. The alternate route may be from a peer, transit provider, or different customer.

Special considerations with regard to the above procedure may be needed for DDoS mitigation service providers. They typically originate or announce a DDoS victim's prefix to their own ISP on a short notice during a DDoS emergency. Some provisions would need to be made for such cases, and they can be determined with the help of inputs from DDoS mitigation service providers.

For developing a list of all the ASes (Cust_AS_List) that are in the customer cone of an ISP, the AS path based Outbound Route Filter (ORF) technique [I-D.ietf-idr-aspath-orf] can be helpful (see discussion in Appendix D.4).

Another technique based on AS_PATH filters is described in [Snijders]. This method is applicable to very large ISPs that have lateral peering. For a pair of such very large ISPs, say A and B, the method depends on ISP A communicating out-of-band (e.g., by email) with ISP B about whether or not it (ISP A) has any transit providers. This out-of-band knowledge enables ISP B to apply suitable AS_PATH filtering criteria for routes involving the presence of ISP A in the path and prevent certain kinds of route leaks (see [Snijders] for details).

Appendix D. Design Rationale and Discussion

This section provides design justifications for the methodology specified in Section 3, and also answers some questions that are anticipated or have been raised in the IETF IDR and SIDR working group meetings.

D.1. Is route-leak solution without cryptographic protection a serious attack vector?

It has been asked if a route-leak solution without BGPsec, i.e., when RLP Fields are not protected, can turn into a serious new attack vector. The answer seems to be: not really! Even the NLRI and AS_PATH in BGP updates are attack vectors, and RPKI/OV/BGPsec seek to fix that. Consider the following. Say, if 99% of route leaks are

accidental and 1% are malicious, and if route-leak solution without BGPsec eliminates the 99%, then perhaps it is worth it (step in the right direction). When BGPsec comes into deployment, the route-leak protection (RLP) bits can be mapped into BGPsec (using the Flags field) and then necessary security will be in place as well (within each BGPsec island as and when they emerge).

Further, let us consider the worst-case damage that can be caused by maliciously manipulating the RLP Field values in an implementation without cryptographic protection (i.e., sans BGPsec). Manipulation of the RLP bits can result in one of two types of attacks: (a) Upgrade attack and (b) Downgrade attack. Descriptions and discussions about these attacks follow. In what follows, P2C stands for transit provider to customer (Down); C2P stands for customer to transit provider (Up), and p2p stands for peer to peer (lateral or non-transit relationship).

(a) Upgrade attack: An AS that wants to intentionally leak a route would alter the RLP encodings for the preceding hops from 1 (i.e., 'Do not Propagate Up or Lateral') to 0 (default) wherever applicable. This poses no problem for a route that keeps propagating in the 'Down' (P2C) direction. However, for a route that propagates 'Up' (C2P) or 'Lateral' (p2p), the worst that can happen is that a route leak goes undetected. That is, a receiving router would not be able to detect the leak for the route in question by the RLP mechanism described here. However, the receiving router may still detect and mitigate it in some cases by applying other means such as prefix filters [RFC7454]. If some malicious leaks go undetected (when RLP is deployed without BGPsec) that is possibly a small price to pay for the ability to detect the bulk of route leaks that are accidental.

(b) Downgrade attack: RLP encoding is set to 1 (i.e., 'Do not Propagate Up or Lateral') when it should be set to 0 (default). This would result in a route being mis-detected and marked as a route leak. By default, RLP encoding is set to 0, and that helps reduce errors of this kind (i.e., accidental downgrade incidents). Every AS or ISP wants reachability for prefixes it originates and for its customer prefixes. So, an AS or ISP is not likely to change an RLP value 0 to 1 intentionally. If a route leak is detected (due to intentional or accidental downgrade) by a receiving router, it would prefer an alternate 'clean' route from a transit provider or peer over a 'marked' route from a customer. It may end up with a suboptimal path. In order to have reachability, the receiving router would accept a 'marked' route if there is no alternative that is 'clean'. So, RLP downgrade attacks (intentional or accidental) would be quite rare, and the consequences do not appear to be grave.

D.2. **Combining results of route-leak detection, OV and BGPsec validation for path selection decision**

Combining the results of route-leak detection, OV, and BGPsec validation for path selection decision is up to local policy in a receiving router. As an example, a router may always give precedence to outcomes of OV and BGPsec validation over that of route-leak detection. That is, if an update fails OV or BGPsec validation, then the update is not considered a candidate for path selection. Instead, an alternate update is chosen that passed OV and BGPsec validation and additionally was not marked as route leak.

If only OV is deployed (and not BGPsec), then there are six possible combinations between OV and route-leak detection outcomes. Because there are three possible outcomes for OV (NotFound, Valid, and Invalid) and two possible outcomes for route-leak detection (marked as leak and not marked). If OV and BGPsec are both deployed, then there are twelve possible combinations between OV, BGPsec validation, and route-leak detection outcomes. As stated earlier, since BGPsec protects the RLP encoding, there would be added certainty in route-leak detection outcome if an update is BGPsec valid (see Appendix D.1).

D.3. **Are there cases when valley-free violations can be considered legitimate?**

There are studies in the literature [Anwar] [Giotsas] [Wijchers] observing and analyzing the behavior of routes announced in BGP updates using data gathered from the Internet. In particular, the studies have focused on how often there appear to be valley-free (e.g., Gao-Rexford [Gao] model) violations, and if they can be explained [Anwar]. One important consideration for explanation of violations is per-prefix routing policies, i.e., routes for prefixes with different peering relationships may be sent over the same link. One encouraging result reported in [Anwar] is that when per-prefix routing policies are taken into consideration in the data analysis, more than 80% of the observed routing decisions fit the valley-free model (see Section 4.3 and SPA-1 data in Figure 2). [Anwar] also observes, "it is well known that this model [the basic Gao-Rexford model and some variations of it] fails to capture many aspects of the interdomain routing system. These aspects include AS relationships that vary based on the geographic region or destination prefix, and traffic engineering via hot-potato routing or load balancing." So, there may be potential for explaining the remaining (20% or less) violations of valley-free as well.

One major design factor is that the Route-Leak Protection (RLP) encoding is per prefix. Hence, the solution is consistent with ISPs'

per-prefix routing policies. Large global and other major ISPs will be the likely early adopters, and they are expected to have expertise in setting policies (including per prefix policies, if applicable), and make proper use of the RLP indications on a per prefix basis. When the large ISPs participate in this solution deployment, it is envisioned that they would form a ring of protection against route leaks, and co-operatively avoid many of the common types of route leaks that are observed. Route leaks may still happen occasionally within the customer cones (if some customer ASes are not participating or not diligently implementing RLP), but such leaks are unlikely to propagate from one large participating ISP to another.

D.4. Comparison with other methods (routing security BCPs)

It is reasonable to ask if techniques considered in BCPs such as[RFC7454] (BGP Operations and Security) and [NIST-800-54] may be adequate to address route leaks. The prefix filtering recommendations in the BCPs may be complementary but not adequate. The difficulty is in ISPs' ability to construct prefix filters that represent their customer cones (CC) accurately, especially when there are many levels in the hierarchy within the CC. In the RLP-encoding based solution described here, AS operators signal for each route propagated, if it must not be subsequently propagated to a transit provider or peer.

AS path based Outbound Route Filter (ORF) described in [I-D.ietf-idr-aspath-orf] is also an interesting complementary technique. It can be used as an automated collaborative messaging system (implemented in BGP) for ISPs to try to develop a complete view of the ASes and AS paths in their CCs. Once an ISP has that view, then AS path filters can be possibly used to detect route leaks. One limitation of this technique is that it cannot duly take into account the fact that routes for prefixes with different peering relationships may be sent over the same link between ASes. Also, the success of AS path based ORF depends on whether ASes at all levels of the hierarchy in a CC participate and provide accurate information (in the ORF messages) about the AS paths they expect to have in their BGP updates.

D.5. Per-Hop RLP Field or Single RLP Flag per Update?

The route-leak detection and mitigation mechanism described in this document is based on setting RLP Fields on a per-hop basis. There is another possible mechanism based on a single RLP flag per update.

Method A - Per-Hop RLP Field: The sender (eBGP router) on each hop in the AS path sets its RLP Field = 1 if sending the update to a customer or lateral peer (see Section 3.2) and Section 3.2.1). No AS

(if operating correctly) would rewrite the RLP Field set by any preceding AS.

Method B - Single RLP Flag per Update: As it propagates, the update would have at most one RLP flag. Once an eBGP router (in the update path) determines that it is sending an update towards a customer or lateral peer AS, it sets the RLP flag. The flag value equals the AS number of the eBGP router that is setting it. Once the flag is set, subsequent ASes in the path must propagate the flag as is.

To compare Methods A and B, consider the example illustrated in Figure 3. Consider a partial deployment scenario in which AS1, AS2, AS3 and AS5 participate in RLP, and AS4 does not. AS1 (2 levels deep in AS3's customer cone) has imperfect RLP operation. Each complying AS's route leak mitigation policy is to prefer an update not marked as route leak (see Section 3.4). If there is no alternative, then a transit-provider may propagate a marked update from a customer. In this example, multi-homed AS4 leaks a route received for prefix Q from transit-provider AS3 to transit-provider AS5.

```
                     +-----------+  RLP=1    +-----------+
                     |    AS3    |---------->|    AS5    |
                     |(Major ISP)|        U2 |(Major ISP)|
                     +-----------+           +-----------+
                        /\       \            /\ U1
 Route for Q propagated /         \RLP=1      /
   due to lack of      /RLP=0      \         / (route leak;
  alternate route     /            \/       /   bad behavior)
              +---------+       +-------------+
              |   AS2   |       |     AS4     |
              +---------+       +-------------+
                  /\             (legacy; does not support RLP)
                  /
                 /
                /RLP=1 (set incorrectly)
               /
     +----------+
     |    AS1   |
     +----------+
        /\
        /
       / Prefix Q
```

Figure 3: Example for comparison of Method A vs. Method B

If Method A is implemented in the network, the two BGP updates for prefix Q received at AS5 are (note that AS4 is not participating in RLP):

U1A: Q [AS4 AS3 AS2 AS1] {RLP3(AS3)=1, RLP2(AS2)=0, RLP1(AS1)=1} ..... from AS4

U2A: Q [AS3 AS2 AS1] {RLP3(AS3)=1, RLP2(AS2)=0, RLP1(AS1)=1} ..... from AS3

Alternatively, if Method B is implemented in the network, the two BGP updates for prefix Q received at AS5 are:

U1B: Q [AS4 AS3 AS2 AS1] {RLP(AS1)=1} ..... from AS4

U2B: Q [AS3 AS2 AS1] {RLP(AS1)=1} ..... from AS3

All received routes for prefix Q at AS5 are marked as route leak in either case (Method A or B). In the case of Method A, AS5 can use additional information gleaned from the RLP fields in the updates to possibly make a better best path selection. For example, AS5 can determine that U1A update received from its customer AS4 exhibits violation of two RLP fields (those set by AS1 and AS3) and one of them was set just two hops away. But U2A update exhibits that only one RLP field was violated and that was set three hops back. Based on this logic, AS5 may prefer U2A over U1A (even though U1A is a customer route). This would be a good decision. However, Method B does not facilitate this kind of more rational decision process. With Method B, both updates U1B and U2B exhibit that they violated only one RLP field (set by AS1 several hops away). AS5 may then prefer U1B over U2B since U1B is from a customer, and that would be bad decision. This illustrates that, due to more information in per-hop RLP Fields, Method A seems to be operationally more beneficial than Method B.

Further, for detection and notification of neighbor AS's non-compliance, Method A (per-hop RLP) is better than Method B (single RLP). With Method A, the bad behavior of AS4 would be explicitly evident to AS5 since it violated AS3's (only two hops away) RLP field as well. AS5 would alert AS4 and also AS2 would alert AS1 about lack of compliance (when Method A is used). With Method B, the alerting process may not be as expeditious.

Acknowledgements

The authors wish to thank Jared Mauch, Jeff Haas, Job Snijders, Warren Kumari, Amogh Dhamdhere, Jakob Heitz, Geoff Huston, Randy Bush, Alexander Azimov, Ruediger Volk, Sue Hares, Wes George, Job

Snijders, Chris Morrow, Sandy Murphy, Danny McPherson, and Eric Osterweil for comments, suggestions, and critique. The authors are also thankful to Padma Krishnaswamy, Oliver Borchert, and Okhee Kim for their review and comments.

Authors' Addresses

Kotikalapudi Sriram
USA National Institute of Standards and Technology

Email: ksriram@nist.gov

Doug Montgomery
USA National Institute of Standards and Technology

Email: dougm@nist.gov

Brian Dickson

Email: brian.peter.dickson@gmail.com

Keyur Patel
Arrcus

Email: keyur@arrcus.com

Andrei Robachevsky
Internet Society

Email: robachevsky@isoc.org