Network Working Group   M. McBride
Internet-Draft   Futurewei
Intended status: Best Current Practice   D. Madory
Expires: February 19, 2021   Oracle
                                                            J. Tantsura
                                                                 Apstra
                                                              R. Raszuk
                                                           Bloomberg LP
                                                                  H. Li
                                                                     HP
                                                        August 18, 2020

                            AS Path Prepending
                  draft-ietf-grow-as-path-prepending-00

Abstract

   AS Path Prepending provides a tool to manipulate the BGP AS_Path
   attribute through prepending multiple entries of an AS. AS Path
   Prepending is used to deprioritize a route or alternate path. By
   prepending the local ASN multiple times, ASes can make advertised AS
   paths appear artificially longer. Excessive AS Path Prepending has
   caused routing issues in the internet. This document provides
   guidance,to the internet community, with how best to utilize AS Path
   Prepending in order to avoid negatively affecting the internet.

Copyright Notice

Table of Contents

1. Introduction

   The Border Gateway Protocol (BGP) [RFC4271] specifies the AS_Path
   attribute which enumerates the ASs that must be traversed to reach
   the networks listed in the BGP UPDATE message. If the UPDATE message
   is propagated over an external link, then the local AS number is
   prepended to the AS_PATH attribute, and the NEXT_HOP attribute is
   updated with an IP address of the router that should be used as a
   next hop to the network. If the UPDATE message is propagated over an
   internal link, then the AS_PATH attribute and the NEXT_HOP attribute
   are passed unmodified.

A common practice among operators is to prepend multiple entries of
an AS (known as AS Path Prepending) in order to deprioritize a route
or a path. This has worked well in practice but the practice is
increasing, with both IPv4 and IPv6, and there are inherit risks to
the global internet especially with excessive AS Path Prepending.
Prepending is frequently employed in an excessive manner such that it
renders routes vulnerable to disruption or misdirection. AS Path
Prepending is discussed in Use of BGP Large Communities [RFC8195] and
this document provides additional, and specific, guidance to
operators on how to be a good internet citizen with the proper use of
AS Path Prepending.

1.1. 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 RFC 2119 [RFC2119].

2. Use Cases

There are various reasons that AS Path Prepending is in use today
including:

o Preferring one ISP over another ISP on the same ASBR or across
   different ASBRs

o Preferring one ASBR over another ASBR in the same site

o Utilize one path exclusively and another path solely as a backup

o Signal to indicate that one path may have a different amount of
   capacity than another where the lower capacity link still takes
   traffic

The following illustration, from Geoff Hustons Path Prepending in BGP
[1], shows how AS Prepending is typically used:

```
                    +---+      +---+
              +---| D |----| F |
              |     +---+      +---+
   +---+    +---+                      |
   | A |---| B |                      |
   +---+    +---+                      |
              |     +---+      +---+
              +---| C |----| E |
                    +---+      +---+
```
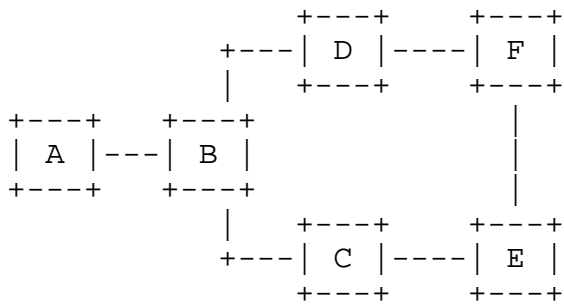
B will normally prefer the path via C to send traffic to E, as this
represents the shorter AS path for B. If E were to prepend a further
two instances of its own AS number when advertising its routes to C,
then B will now see a different situation, where the AS Path via D
represents the shorter path. Through the use of selective prepending
E is able to alter the routing decision of B, even though B is not an
adjacent neighbour of E. The result is that traffic from A and B
will be passed via D and F to reach E, rather than via C. In this
way prepending implements action at a distance where the routing
decisions made by non-adjacent AS's can be influenced by selective AS
Path prepending.

3. Problems

   Since it it so commonly used, what is the problem with the excessive
   use of AS Path Prepending? Here are a few examples:

3.1. Excessive Prepending

   The risk of excessive use of AS Path Prepending can be illustrated
   with real-world examples that have been anonymized using documention
   prefixes [RFC5737] and AS's [RFC5398] . Consider the prefix
   198.51.100.0/24 which is normally announced with an inordinate amount
   of prepending. A recent analysis revealed that 198.51.100.0/24 is
   announced to the world along the following AS path:

   64496 64511 64511 64511 64511 64511 64511 64511 64511 64511 64511
   64511 64511 64511 64511 64511 64511 64511 64511 64511 64511 64511
   64511 64511

   In this example, the origin AS64511 appears 23 consecutive times
   before being passed on to a single upstream (AS64496), which passes
   it on to the global internet, prepended-to-all. An attacker, wanting
   to intercept or manipulate traffic to this prefix, could enlist a
   datacenter to allow announcements of the same prefix with a
   fabricated AS path such as 999999 64496 64511. Here the fictional

AS999999 represents the shady datacenter. This malicious route would
be preferred due to the shortened AS path length and might go
unnoticed by the true origin, even if route-monitoring had been
implemented. Standard BGP route monitoring checks a route's origin
and upstream and both would be intact in this scenario. The length
of the prepending gives the attacker room to craft an AS path that
would appear plausible to the casual observer, comply with origin
validation mechanisms, and not be detected by off-the-shelf route
monitoring.

3.2. Prepending during a routing leak

In April 2010, a service provider experienced a routing leak. While
analyzing the leak something peculiar was noticed. When we ranked
the approximately 50,000 prefixes involved in the leak based on how
many ASes accepted the leaked routes, most of the impact was
constrained to Country A routes. However, two of the top five most-
propagated leaked routes (listed in the table below) were Country B
routes.

During the routing leak, nearly all of the ASes of the internet
preferred the Country A leaked routes for 192.0.2.0/21 and
198.51.100.0/22 because, at the time, these two Country B prefixes
were being announced to the entire internet along the following
excessively prepended AS path: 64496 64500 64511 64511 64511 64511
64511 64511. Virtually any illegitimate route would be preferred
over the legitimate route. In this case, the victim is all but
ensuring their victimhood.

There was only a single upstream seen in the prepending example from
above, so the prepending was achieving nothing except incurring risk.
You would think such mistakes would be relatively rare, especially
now, 10 years later. As it turns out, there is quite a lot of
prepending-to-all going on right now and during leaks, it doesn't go
well for those who make this mistake. While one can debate the
merits of prepending to a subset of multiple transit providers, it is
difficult to see the utility in prepending to every provider. In
this configuration, the prepending is no longer shaping route
propagation. It is simply incentivizing ASes to choose another
origin if one were to suddenly appear whether by mistake or
otherwise.

3.3. Prepending to All

Based on analysis done in 2019, Excessive AS Path Prepending [2], out
of approximately 750,000 routes in the IPv4 global routing table,
nearly 60,000 BGP routes are prepended to 95% or more of hundreds of
BGP sources. About 8% of the global routing table, or 1 out of every

12 BGP routes, is configured with prepends to virtually the entire
internet. The 60,000 routes include entities of every stripe:
governments, financial institutions, even important parts of internet
infrastructure.

Much of the worst propagation of leaked routes during big leak events
have been due to routes being prepended-to-all. AS64505 leak of
April 2014 (>320,000 prefixes) was prepended-to-all. And the AS64506
leak of June 2015 (>260,000 prefixes) was also prepended-to-all.
Prepended-to-all prefixes are those seen as prepended by all (or
nearly all) of the ASes of the internet. In this configuration,
prepending is no longer shaping route propagation but is simply
incentivizing ASes to choose another origin if one were to suddenly
appear whether by mistake or otherwise. The percentage of the IPv4
table that is prepended-to-all is growing at 0.5% per year. The IPv6
table is growing slower at 0.2% per year. The reasons for using
prepend-to-all appears to be due to 1) the AS forgetting to remove
the prepending for one of its transit providers when it is no longer
needed and 2) the AS attempting to de-prioritize traffic from transit
providers over settlement-free peers and 3) there are simply a lot of
errors in BGP routing. Consider the prepended AS path below:

64496 64501 64501 64510 64510 64501 64510 64501 64501 64510 64510
64501 64501 64510

The prepending here involves a mix of two distinct ASNs (64501 and
64510) with the last two digits transposed.

3.4. Memory

BGP attribute sets are shared among stored routes, ie, if two stored
routes have the same attribute sets, the attribute set is stored only
once. AS Paths are shared among attribute sets so that if two stored
attribute sets have the same AS Path, then the AS Path is stored only
once. Storing them in the control plane is not a big problem.
However, AS Paths can be sent in Netflow which is generated in the
forwarding plane. AS Paths are not stored in expensive fast memory
on the forwarding plane, but still, using memory on the forwarding
plane has greater impact than on the control plane. An AS Path
consists of AS_SEQUENCE (and other elements). An AS_SEQUENCE can
contain a maximum of 255 ASNs. If the AS Path is longer, then
multiple AS_SEQUENCE's are required. The code to parse them and
create them is not often exercised and is a potential for bugs in
fresh code. The older implementations have these bugs well and truly
shaken out of them. Some BGP implementations have had memory
corruption/fragmentation problems with long AS Paths.

3.5. Errant announcement

   There was an Internet-wide outage caused by a single errant routing
   announcement. In this incident, AS64496 announced its one prefix
   with an extremely long AS path. Someone entered their ASN instead of
   the prepend count 64496 modulo 256 = 252 prepends and when a path
   lengths exceeded 255, routers crashed

3.6. Alternatives to AS Path Prepend

   Robert: Let me try to put together various options on how could we
   provide the above objective without trashing Internet with long AS-
   PATHs. I can think of at least few options. Note that since some
   vendors ship you the tools to modify any part of the AS-PATH while in
   transit or for that matter BGP code is commodity and one could modify
   anything in the BGP UPDATE no matter what we recommend or even
   standardize there is no assurance that it will always work.

4. Best Practices

   Many of the best practices, or lack thereof, can be illustrated from
   the preceeding examples. Here's a summary of the best current
   practices of using AS Path Prepending:

   Doug ToDo: histogram or table of the frequency of unique-aspath
   lengths (i.e. with all prepending flattened). This would help
   contextualise the best practice recommendation.

   o Network operators should ensure prepending is absolutely
      necessary. Many of your networks have excessive prepending

   o Prepending more than 5 times buys you nothing. So don't do it.

   o Prepending-to-all is a self-inflicted and needless risk that
      serves little purpose. Those excessively prepending their routes
      should consider this risk and adjust their routing configuration.

   o It is not typical to see more than 20 ASes in a AS_PATH in the
      Internet today even with the use of AS_Path prepend. The Internet
      is typically around 5 ASes deep with the largest AS_PATH being
      16-20 ASNs. Some have added 100 or more AS Path Prepends and
      operators should therefore consider limiting the maximum AS-path
      length being accepted

5. IANA Considerations

6. Security Considerations

   There are no security issues introduced by this draft.

7. Acknowledgement

   The authors would like to thank Greg Skinner, Randy Bush, Dave
   Farmer, Nick Hilliard, Martijn Schmidt, Jakob Heitz, Michael Still
   and Geoff Huston for contributing to this document.

8. References

8.1. Normative References

   [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate
              Requirement Levels", BCP 14, RFC 2119,
              DOI 10.17487/RFC2119, March 1997,
              <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>.

   [RFC5398] Huston, G., "Autonomous System (AS) Number Reservation for
              Documentation Use", RFC 5398, DOI 10.17487/RFC5398,
              December 2008, <https://www.rfc-editor.org/info/rfc5398>.

   [RFC5737] Arkko, J., Cotton, M., and L. Vegoda, "IPv4 Address Blocks
              Reserved for Documentation", RFC 5737,
              DOI 10.17487/RFC5737, January 2010,
              <https://www.rfc-editor.org/info/rfc5737>.

   [RFC8195] Snijders, J., Heasley, J., and M. Schmidt, "Use of BGP
              Large Communities", RFC 8195, DOI 10.17487/RFC8195, June
              2017, <https://www.rfc-editor.org/info/rfc8195>.

8.2. URIs

   [1] https://labs.apnic.net/?p=1264

   [2] https://blogs.oracle.com/internetintelligence/excessive-as-path-
       prepending-is-a-self-inflicted-vulnerability

Authors' Addresses

    Mike McBride
    Futurewei

    Email: michael.mcbride@futurewei.com

    Doug Madory
    Oracle

    Email: douglas.madory@oracle.com

    Jeff Tantsura
    Apstra

    Email: jefftant.ietf@gmail.com

    Robert Raszuk
    Bloomberg LP

    Email: robert@raszuk.net

    Hongwei Li
    HP

    Email: flycoolman@gmail.com