

Network Working Group
Internet-Draft
Intended status: Best Current Practice
Expires: 10 August 2024

M. McBride
Futurewei
D. Madory
Kentik
J. Tantsura
Nvidia
R. Raszuk
NTT Network Innovations
H. Li
HPE
J. Heitz
Cisco
G. Mishra
Verizon Inc.
7 February 2024

AS Path Prepending
draft-ietf-grow-as-path-prepend-12

Abstract

AS Path Prepending provides a tool to manipulate the BGP AS_PATH attribute through prepending multiple entries of an ASN. AS Path Prepending is used to deprioritize a route or alternate path. By prepending the local ASN multiple times, ASs can make advertised AS paths appear artificially longer. Excessive AS Path Prepending has caused routing issues in the Internet. This document provides guidance for the use of AS Path Prepending, including alternative solutions, in order to avoid negatively affecting the Internet.

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 10 August 2024.

Copyright Notice

Copyright (c) 2024 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 Revised BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Revised BSD License.

Table of Contents

1. Introduction	2
1.1. Requirements Language	3
2. Use Cases	3
3. Problems	4
3.1. Cascading and ripple effects of prepending across the Internet	4
3.2. Excessive Prepending	5
3.3. Prepending during a routing leak	6
3.4. Prepending to All	7
3.5. Memory	7
3.6. Errant announcement	8
4. Alternatives to AS Path Prepend	8
5. Best Practices	9
6. IANA Considerations	11
7. Security Considerations	11
8. Acknowledgement	11
9. References	11
9.1. Normative References	11
9.2. Informative References	12
Authors' Addresses	12

1. Introduction

The Border Gateway Protocol (BGP) [RFC4271] specifies the AS_PATH attribute, which enumerates ASs a route update has traversed. 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. So far, this has not caused many problems. However, the practice is increasing, with both IPv4 and IPv6, and there are now inherent 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. [RFC8195] discusses using BGP Large Communities for traffic engineering through selective AS_PATH prepending. This document provides additional and specific guidance to operators on how to be good Internet citizens with less risky use of AS Path Prepending.

1.1. Requirements Language

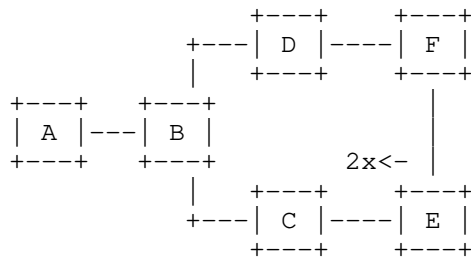
The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

2. Use Cases

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

- * Preferring one ISP over another ISP on the same ASBR or across different ASBRs.
- * Preferring one ASBR over another ASBR in the same site or between sister sites.
- * Utilize one path exclusively and another path solely as a backup.
- * Signal to indicate that one path may have a different amount of capacity than another where the lower capacity link still takes traffic.
- * Conditionally prefer one ASBR over another at the same site or between sites for lowest latency path based on geographic location.
- * An ISP doesn't accept traffic engineering using BGP communities. Prepending is the only option.

The following illustration, from Geoff Huston's [Path_Prepending_in_BGP], shows how AS Prepending is typically used:



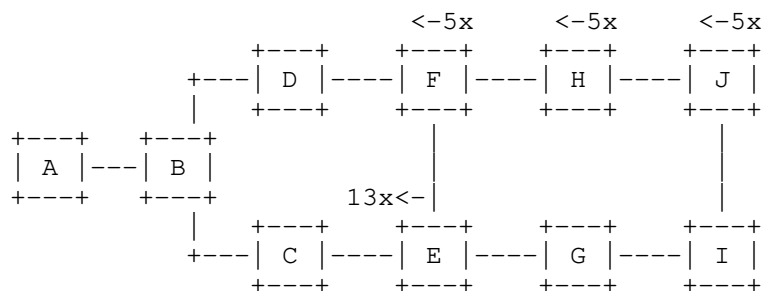
In the diagram above, A, B, C, D, E, and F all have a different AS. B will normally prefer the path via C to send traffic to E, as this represents the shorter AS path for B. If E were to prepend a further two instances of its own AS number when advertising its routes to C, then B will now see a different 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 ASs can be influenced by selective AS Path prepending.

3. Problems

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

3.1. Cascading and ripple effects of prepending across the Internet

Care should be taken in prepending, as prepending can cause ripple effects with multiple AS's performing the same set of prepends in the same direction, resulting in unintended routing where the valid preferred path becomes now de-preferred.



In the diagram above A, B, C, D, E, F, G, H, I, and J are all part of different ASes. B will normally prefer the path via D to send traffic to J, as this represents the preferred path to B, due to E prepending 13 instances of its own AS number when advertising routes to C. ISP J decides to prepend 5 instances of its own AS when advertising to H, and ISP H decides to do the same and prepends 5 instances of its own AS when advertising to F. ISP F decides as well to prepend 5 instances of its own AS when advertising to D. B now sees 19 AS hops for prefixes coming from D to get to J which should be the preferred path compared to 18 AS hops coming from C which is now preferred. We now have a route leak to I as B now sends all of its traffic through I to reach J. This is the typical scenario where route leaks occur where providers decide to de-prefer a path. However as the same de-preference of a path gets cascaded in the same direction, as a result, the path that should never be preferred as its as-path is very high in this case 18 AS hops ends up being the preferred path resulting in a route leak. Usage of BGP large communities along with conditional prepending, along with care being taken when prepending is performed between providers, can help mitigate the adverse impacts of prepending.

3.2. Excessive Prepending

The risk of excessive use of AS Path Prepending can be illustrated with real-world examples that have been anonymized using documentation prefixes [RFC5737] and ASs [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 routes 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.

3.3. 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 ASs 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 ASs 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 doesnt 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 ASs to choose another origin if one were to suddenly appear whether by mistake or otherwise.

3.4. Prepending to All

Based on analysis done in 2019 [Excessive_AS_Path_Prepending], 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. The 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. Prepend-to-all prefixes are those seen as prepended by all (or nearly all) of the ASs of the Internet. In this configuration, prepending is no longer shaping route propagation but is simply incentivizing ASs 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.

3.5. Memory

Long AS Paths cause an increase in memory usage by BGP speakers. A concern about an AS Path longer than 255 is the extra complexity required to process it, because it needs to be encoded in more than one AS_SEQUENCE in the AS_PATH BGP path attribute.

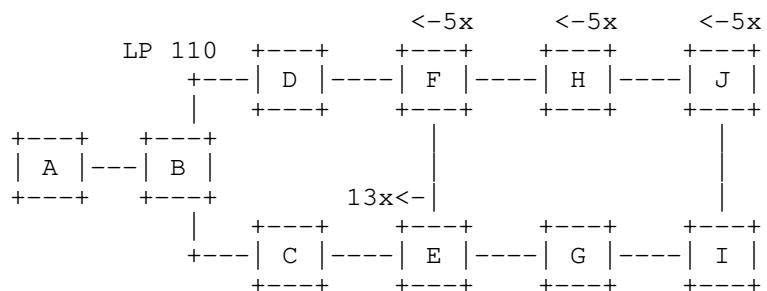
3.6. Errant announcement

It is possible for an Internet-wide outage to occur because of a single errant routing announcement. For example, AS64496 could announce its one prefix with an extremely long AS path. Someone could enter their ASN instead of the prepend count. $64496 \bmod 256 = 240$ prepends, and when a path length exceeded 255, routers could crash.

4. Alternatives to AS Path Prepend

Various options, to provide path preference without needing to use AS Path Prepend, include:

- * Use predefined communities that are mapped to a particular behavior when propagated.
- * Announce more specific routes on the preferred path.
- * The BGP Origin Code is an attribute that is used for path selection and can be used as a high order tie-breaker. The three origin codes are IGP, EGP and INCOMPLETE. When AS Paths are of equivalent length, users could advertise paths, with IGP or EGP origin, over the preferred path while the other ASBRs (which would otherwise need to prepend N times) advertises with an INCOMPLETE origin code.
- * The Multi Exit Discriminator (MED) is an optional non-transitive attribute that can be used to influence path preference instead of using as-path. MED is non transitive so it cannot influence an AS more than 1 AS hop away.
- * Local-preference optional non-transitive attribute, above as-path in BGP path selection, can be used to influence route preference within the local operators AS administrative domain. Local-preference can shield the operator domain from traffic shifts off the preferred path to a de-preferred path caused by excess prepending done by service providers across the Internet. If all service providers across the Internet set local-preference inbound conditionally with Large Community set on preferred paths, the impacts of suboptimal routing, as well as other routing issues resulting from excess prepending, can be mitigated.



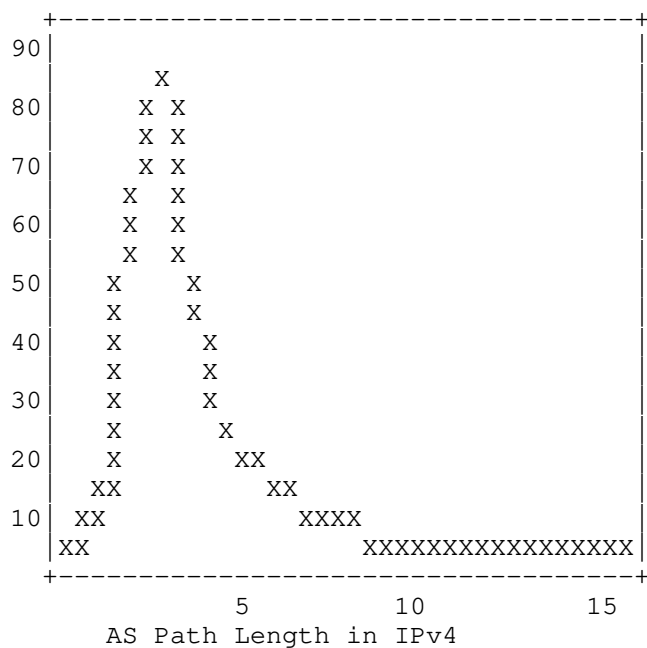
In the diagram above A, B, C, D, E, F, G, H, I, J are all part of a different AS. B will normally prefer the path via D to send traffic to J, as this represents the preferred path to B, due to E prepending 13 instances of its own AS number when advertising routes to C. ISP J decides to prepend 5 instances of its own AS when advertising to H, and ISP H decides to do the same and prepends 5 instances of its own AS when advertising to F. ISP F decides to also prepend 5 instances of its own AS when advertising to D. B now sees 19 AS hops for prefixes coming from D to get to J which should be the preferred path compare to 18 AS hops coming from C which is now preferred. We now have suboptimal routing to I as B now sends all of its traffic through I to reach J. This suboptimal routing on B can be prevented locally within the operator domain by setting local-preference inbound, which is above as-path length in the best path selection, to higher than default 100 to 110 inbound from D, thus shielding the operator B from being influenced by the excessive prepend cascading ripple affect by F, H, J. Note that A still sees the cascading ripple affect of excess prepending, however A, or any service provider AS downstream of B, ingressing B, will be shunted to D via local-preference and the suboptimal routing is now mitigated for all downstream AS to the left of B that prefer the path through B.

5. Best Practices

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

- * Network operators should ensure prepending is absolutely necessary as many networks have excessive prepending. It is best to innumerate what the routing policies are intended to achieve before concluding that prepending is a solution

- * The neighbor you are prepending may have an unconditional preference for customer routes and prepending doesn't work. It's helpful to check with neighbors to see if they will honor the prepend to avoid wasting effort and potentially causing further vulnerabilities.
- * Use of local-preference inbound on preferred paths between service providers to help mitigate the adverse affects of prepending
- * There is no need to prepend more than 5 ASs. The following diagram, from the previously referenced AS Path Prepending analysis from 2019, shows that 90% of AS path lengths are 5 ASNs or fewer in length.



X Axis = unique AS Paths in millions

- * Don't prepend ASNs that you don't own.
- * Don't prepend if your AS is single homed.

- * 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.
- * The Internet is typically around 5 ASs 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 through aggressive filter policies.

6. IANA Considerations

7. Security Considerations

Long prepending may make a network more vulnerable to route hijacking which will exist whenever there is a well connected peer that is willing to forge their AS_PATH or allows announcements with a fabricated AS path. Accepting routes with extremely long AS_PATHs may cause increased memory usage and possible router crashes. Using Autonomous System Provider Authorization (ASPA) objects in the Resource Public Key Infrastructure (RPKI), to verify the BGP AS_PATH attribute of advertised routes, would provide detection and mitigation of route leaks and improbable AS paths.

For a more comprehensive discussion of BGP Operations and Security, see [RFC7454].

8. Acknowledgement

The authors would like to thank Greg Skinner, Randy Bush, David Farmer, Nick Hilliard, Martijn Schmidt, Michael Still, Geoff Huston, Jeffrey Haas, Alejandro Acosta and Martin Pels for contributing to this document.

9. References

9.1. 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>>.
- [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>>.

9.2. Informative References

- [Excessive_AS_Path_Prepending]
Madory, D., "Excessive AS Path Prepending", Article APNIC, 2019, <<https://blog.apnic.net/2019/07/15/excessive-bgp-as-path-prepend-is-a-self-inflicted-vulnerability/>>.
- [Path_Prepending_in_BGP]
Huston, J., "Path Prepending in BGP", Article APNIC, 2019, <<https://labs.apnic.net/index.php/2019/10/27/path-prepend-in-bgp/>>.
- [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>>.

Authors' Addresses

Mike McBride
Futurewei
Email: michael.mcbride@futurewei.com

Doug Madory
Kentik
Email: dmadory@kentik.com

Jeff Tantsura
Nvidia

Email: jefftant.ietf@gmail.com

Robert Raszuk
NTT Network Innovations
940 Stewart Dr
Sunnyvale, CA 94085
United States of America
Email: robert@raszuk.net

Hongwei Li
HPE
Email: flycoolman@gmail.com

Jakob Heitz
Cisco
170 West Tasman Drive
San Jose, CA 95134
United States of America
Email: jheitz@cisco.com

Gyan Mishra
Verizon Inc.
Email: gyan.s.mishra@verizon.com