Workgroup: Network Working Group
Internet-Draft:
draft-ietf-idr-deprecate-as-set-confed-set-12
Obsoletes: 6472 (if approved)
Updates: 4271 5065 (if approved)
Published: 10 January 2024
Intended Status: Standards Track
Expires: 13 July 2024
Authors: W. Kumari
Google, Inc.
K. Sriram
USA NIST
L. Hannachi
USA NIST
J. Haas
Juniper Networks, Inc.

Deprecation of AS_SET and AS_CONFED_SET in BGP

Abstract

BCP 172 (i.e., RFC 6472) recommends not using AS_SET and AS_CONFED_SET AS_PATH segment types in the Border Gateway Protocol (BGP). This document advances that recommendation to a standards requirement in BGP; it proscribes the use of the AS_SET and AS_CONFED_SET path segment types in the AS_PATH. This is done to simplify the design and implementation of BGP and to make the semantics of the originator of a BGP route clearer. This will also simplify the design, implementation, and deployment of various BGP security mechanisms. This document updates RFC 4271 and RFC 5065, and obsoletes RFC 6472.

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 13 July 2024.

Copyright Notice

Table of Contents

1. Introduction

BCP 172 [RFC6472] makes a recommendation for not using AS_SET (see [RFC4271]) and AS_CONFED_SET (see [RFC5065]) AS_PATH path segment types in the Border Gateway Protocol (BGP). This document advances the BCP recommendation to a standards requirement in BGP; it proscribes the use of the AS_SET and AS_CONFED_SET types of path segments in the AS_PATH. The purpose is to simplify the design and implementation of BGP and to make the semantics of the originator of

a BGP route clearer. This will also simplify the design, implementation, and deployment of various BGP security mechanisms. In particular, the proscription of AS_SETs and AS_CONFED_SETs removes the possibility of ambiguity about origin AS in RPKI-based route origin validation (RPKI-ROV) [RFC6811] [RFC6907] [RFC9319].

The AS_SET path segment in the AS_PATH attribute (Sections 4.3 and 5.1.2 of [RFC4271]) is created by a router that is performing route aggregation and contains an unordered set of Autonomous Systems (ASes) that contributing prefixes in the aggregate have traversed.

The AS_CONFED_SET path segment (see [RFC5065]) in the AS_PATH attribute is created by a router that is performing route aggregation and contains an unordered set of Member AS Numbers in the local confederation that contributing prefixes in the aggregate have traversed. It is very similar to an AS_SET but is used within a confederation.

By performing aggregation, a router is combining multiple BGP routes for more specific destinations into a new route for a less specific destination ([RFC4271], Section 9.1.2.2.). Aggregation may blur the semantics of the origin AS for the prefix being announced by producing an AS_SET or AS_CONFED_SET. Such sets can cause operational issues, such as not being able to authenticate a route origin for the aggregate prefix in new BGP security technologies such as those that take advantage of X.509 extensions for IP addresses and AS identifiers ([RFC3779], [RFC6480], [RFC6811], [RFC6907], [RFC8205], [RFC9319]). This could result in reachability problems for the destinations covered by the aggregated prefix.

From analysis of historical Internet routing data, it is apparent that aggregation that involves AS_SETs is very seldom used in practice on the public Internet (see [Analysis]). When it is used, it is often used incorrectly; only a single AS in the AS_SET is the most common case [Analysis]. Also, very often the same AS appears in the AS_SEQUENCE and the AS_SET in the BGP update. The occurrence of reserved AS numbers ([IANA-SP-ASN]) is also somewhat frequent.

2. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

3. Recommendations

BGP speakers conforming to this document (i.e., conformant BGP speakers) SHOULD NOT locally generate BGP UPDATE messages containing

AS_SETs or AS_CONFED_SETs. Conformant BGP speakers SHOULD NOT send BGP UPDATE messages containing AS_SETs or AS_CONFED_SETs. Upon receipt of such messages, conformant BGP speakers SHOULD use the "treat-as-withdraw" error handling behavior as per [RFC7606].

The document uses normative language such as "SHOULD NOT send" rather than "MUST NOT send" with the intention of allowing some transition time for existing implementations and avoiding abrupt disruptions for the operators currently using AS_SETs or AS_CONFED_SETs. However, it is strongly urged that operators stop sending UPDATEs with AS_SETs or AS_CONFED_SETs as quickly as possible to avoid having UPDATEs dropped by BGP security mechanisms such as RPKI-ROV and BGPsec.

If a network operator wishes to consider BGP UPDATE messages with AS_SETs or AS_CONFED_SETs received from an external BGP peers, they MAY have a feature (knob) in their implementation to do so on a per-peer basis. The operator should understand the full implications of choosing this option.

Network operators SHOULD NOT locally generate any new announcements containing AS_SETs or AS_CONFED_SETs.

BGP security technologies (such as those that take advantage of X.509 extensions for IP addresses and AS identifiers ([RFC3779], [RFC6480], [RFC6811], [RFC8205]) might not support routes with AS_SETs or AS_CONFED_SETs in them. Routes with AS_SETs have no possibility of ever being considered RPKI-ROV valid [RFC6811] [RFC6907].

4. Updates to Existing RFCs

This document deprecates the origination of BGP routes with AS_SET (type 1) (see [RFC4271], Section 4.3).

This document also deprecates the origination of BGP routes with AS_CONFED_SET (type 4) AS_PATH segments (see [RFC5065], Section 3).

BGP speakers conforming to this document — i.e., conformant BGP speakers — SHOULD NOT originate BGP UPDATE messages containing AS_SETs or AS_CONFED_SETs. Upon receipt of BGP routes containing AS_SETs, conformant BGP speakers SHOULD use the "treat-as-withdraw" error handling behavior as per [RFC7606].

4.1. BGP AS_PATH "Brief" Aggregation

Sections 9.1.4 and 9.2.2.2 of [RFC4271] describe BGP aggregation procedures. Appendix F.6 in [RFC4271] describes a generally unimplemented "Complex AS_PATH Aggregation" procedure.

[RFC4271], Section 5.1.6, describing the ATOMIC_AGGREGATE Path Attribute, notes that:

> When a BGP speaker aggregates several routes for the purpose of advertisement to a particular peer, the AS_PATH of the aggregated route normally includes an AS_SET formed from the set of ASes from which the aggregate was formed. In many cases, the network administrator can determine if the aggregate can safely be advertised without the AS_SET, and without forming route loops.
>
> If an aggregate excludes at least some of the AS numbers present in the AS_PATH of the routes that are aggregated as a result of dropping the AS_SET, the aggregated route, when advertised to the peer, SHOULD include the ATOMIC_AGGREGATE attribute.

When BGP AS_PATH aggregation is done according to the [RFC4271], Section 9.2.2.2, procedures and any resulting AS_SETs are discarded, this is typically referred to as "brief" aggregation in implementations. Brief aggregation results in an AS_PATH that has the property (from [RFC4271], Section 9.2.2.2):

> determine the longest leading sequence of tuples (as defined above) common to all the AS_PATH attributes of the routes to be aggregated. Make this sequence the leading sequence of the aggregated AS_PATH attribute.

The ATOMIC_AGGREGATE Path Attribute is subsequently attached to the BGP route, if AS_SETs are dropped.

4.2. Issues with "Brief" AS_PATH Aggregation and RPKI-ROV

While brief AS_PATH aggregation has the desirable property of not containing AS_SETs, the resulting aggregated AS_PATH may contain an unpredictable origin AS. Such an unpredictable origin ASes may result in RPKI-ROV validation issues:

* Depending on the contributing routes to the aggregate route, the resulting origin AS may vary.

* The presence of expected contributing routes may be unpredictable due to route availability from BGP neighbors.

* In the presence of such varying origin ASes, it would be necessary for the resource holder to register Route Origin Authorizations (ROAs) [RFC6482] for each potential origin AS that may result from the expected aggregated AS_PATHs.

4.3. Recommendations to Mitigate Unpredictable AS_PATH origins for RPKI-ROV Purposes

In order to ensure a consistent BGP origin AS is announced for aggregate BGP routes for implementations of "brief" BGP aggregation, the implementation should be configured to truncate the AS_PATH after the right-most instance of the desired origin AS for the aggregate. The desired origin AS could be the aggregating AS itself.

If the resulting AS_PATH would be truncated from the otherwise expected result of BGP AS_PATH aggregation (an AS_SET would not be generated, and/or ASes are removed from the "longest leading sequence" of ASes), the ATOMIC_AGGREGATE Path Attribute SHALL be attached. This is consistent with the intent of Section 5.1.6 of [RFC4271].

5. Operational Considerations

When aggregating prefixes, network operators MUST use brief aggregation. In brief aggregation, the AGGREGATOR and ATOMIC_AGGREGATE Path Attributes are included, but the AS_PATH does not have AS_SET or AS_CONFED_SET path segment types. See Appendix B for examples of brief aggregation while keeping the origin AS unambiguous and generating appropriate ROAs.

When doing the above, operators MUST form the aggregate at the border in the outbound BGP policy and omit any prefixes from the AS that the aggregate is being advertised to. In other words, an aggregate prefix MUST NOT be announced to the contributing ASes. Instead, more specific prefixes (from the aggregate) MUST be announced to each contributing AS, excluding any that were learned from the contributing AS in consideration. See Appendix A for an example of this filtering policy.

Operators MUST install egress filters to block data packets when the destination address belongs to an internal prefix. Similarly, any known single-homed customer prefix MUST also be included in the egress filters except on the interface for that customer. These safeguards mitigate looping in the data plane when connection to such an internal or customer prefix is lost. This mechanism effectively compensates for the lack of the additional loop detection capability accorded by AS_SETs (if they were allowed).

6. Security Considerations

This document deprecates the use of aggregation techniques that create AS_SETs or AS_CONFED_SETs. Obsoleting these path segment types from BGP and removal of the related code from implementations would potentially decrease the attack surface for BGP. Deployments

of new BGP security technologies ([RFC6480], [RFC6811], [RFC8205]) benefit greatly if AS_SETs and AS_CONFED_SETs are not used in BGP.

7. IANA Considerations

This document requires no IANA actions.

8. Acknowledgements

The authors would like to thank John Heasley, Job Snijders, Jared Mauch, Jakob Heitz, Keyur Patel, Douglas Montgomery, Randy Bush, Susan Hares, John Scudder, Curtis Villamizar, Danny McPherson, Chris Morrow, Tom Petch, Ilya Varlashkin, Enke Chen, Tony Li, Florian Weimer, John Leslie, Paul Jakma, Rob Austein, Russ Housley, Sandra Murphy, Steve Bellovin, Steve Kent, Steve Padgett, Alfred Hoenes, and Alvaro Retana for comments and suggestions.

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>.

[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>.

9.2. Informative References

[Analysis] Hannachi, L. and K. Sriram, "Detailed analysis of AS_SETs in BGP updates", NIST Robust Inter-domain Routing Project Website , October 2019, <https://github.com/ksriram25/IETF/blob/main/Detailed-AS_SET-analysis.txt>.

[IANA-SP-ASN] "Special-Purpose Autonomous System (AS) Numbers", <https://www.iana.org/assignments/iana-as-numbers-special-registry/iana-as-numbers-special-registry.xhtml>.

[RFC3779] Lynn, C., Kent, S., and K. Seo, "X.509 Extensions for IP Addresses and AS Identifiers", RFC 3779, DOI 10.17487/

RFC3779, June 2004, <https://www.rfc-editor.org/info/rfc3779>.

[RFC6472] Kumari, W. and K. Sriram, "Recommendation for Not Using AS_SET and AS_CONFED_SET in BGP", BCP 172, RFC 6472, DOI 10.17487/RFC6472, December 2011, <https://www.rfc-editor.org/info/rfc6472>.

[RFC6480] Lepinski, M. and S. Kent, "An Infrastructure to Support Secure Internet Routing", RFC 6480, DOI 10.17487/RFC6480, February 2012, <https://www.rfc-editor.org/info/rfc6480>.

[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>.

[RFC6907] Manderson, T., Sriram, K., and R. White, "Use Cases and Interpretations of Resource Public Key Infrastructure (RPKI) Objects for Issuers and Relying Parties", RFC 6907, DOI 10.17487/RFC6907, March 2013, <https://www.rfc-editor.org/info/rfc6907>.

[RFC7606] Chen, E., Ed., Scudder, J., Ed., Mohapatra, P., and K. Patel, "Revised Error Handling for BGP UPDATE Messages", RFC 7606, DOI 10.17487/RFC7606, August 2015, <https://www.rfc-editor.org/info/rfc7606>.

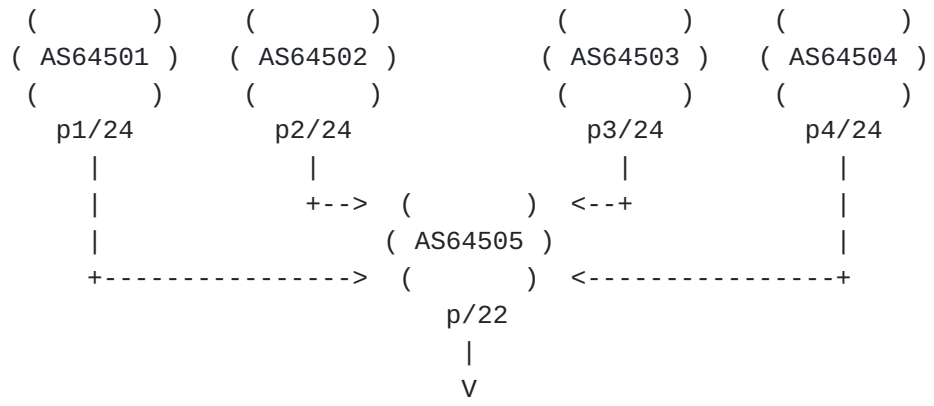
[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <https://www.rfc-editor.org/info/rfc8174>.

[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>.

[RFC9319] Gilad, Y., Goldberg, S., Sriram, K., Snijders, J., and B. Maddison, "The Use of maxLength in the Resource Public Key Infrastructure (RPKI)", BCP 185, RFC 9319, DOI 10.17487/RFC9319, October 2022, <https://www.rfc-editor.org/info/rfc9319>.

Appendix A. Example of Route Filtering for Aggregate Routes and its Contributors

Presented here is an illustration of how an AS_SET is not used when aggregating and still data-plane route loops are avoided. Consider that p1/24 (from AS 64501), p2/24 (from AS 64502), p3/24 (from AS 64503), and p4/24 (from AS 64504) are aggregated by AS 64505 to p/22. AS_SET is not used with the aggregate p/22 but AGGREGATOR and ATOMIC AGGREGATE are used. Data-plane route loops are avoided by not announcing the aggregate p/22 to the contributing ASes, i.e., AS 64501, AS 64502, AS 64503, and AS 64504. Instead, as further illustration, p1/24, p2/24, and p4/24 are announced to AS 64503. The routing tables (post aggregation) of each of the ASes are depicted in the diagram below .



AS 64501
==========
p1/24 AS_PATH ""
p2/24 AS_PATH "64505 64502"
p3/24 AS_PATH "64505 64503"
p4/24 AS_PATH "64505 64504"

AS 64502
==========
p1/24 AS_PATH "64505 64501"
p2/24 AS_PATH ""
p3/24 AS_PATH "64505 64503"
p4/24 AS_PATH "64505 64504"

AS 64503
==========
p1/24 AS_PATH "64505 64501"
p2/24 AS_PATH "64505 64502"
p3/24 AS_PATH ""
p4/24 AS_PATH "64505 64504"

AS 64504
==========
p1/24 AS_PATH "64505 64501"
p2/24 AS_PATH "64505 64502"
p3/24 AS_PATH "64505 64503"
p4/24 AS_PATH ""

AS 64505
==========
p/22 AS_PATH "" AGGREGATOR 64505 ATOMIC_AGGREGATE
p1/24 AS_PATH "64501"
p2/24 AS_PATH "64502"
p3/24 AS_PATH "64503"
p4/24 AS_PATH "64504"

Appendix B. Examples of Inconsistent BGP Origin-AS Generated by Traditional Brief Aggregation

In the examples below, it is illustrated how brief aggregation may result in inconsistent origin AS.

AS 64500 aggregates more specific routes into 192.0.2.0/24.

Consider the following scenarios where brief aggregation is done by AS 64500 and what the resultant origin ASes would be.

Routes:
R1 - 192.0.2.0/26   AS_PATH "64501"
R2 - 192.0.2.64/26 AS_PATH "64502"
R3 - 192.0.2.128/26 AS_PATH "64504 64502"
R4 - 192.0.2.192/26 AS_PATH "64504 64501"

B.1. Scenario 1: First one route, then another, each with a fully disjoint AS_PATH

Receive R1. Aggregate 192.0.2.0/24 AS_PATH "64501"

Alternate "bug?": Aggregate 192.0.2.0/24 AS_PATH "[64501]"

Receive R2. Aggregate 192.0.2.0/24 AS_PATH "[64501 64502]"

If brief aggregation is in use, the AS_PATH would be truncated to the empty AS_PATH, "".

The resulting AS_PATH is thus not stable and depends on the presence of specific routes.

B.2. Scenario 2: First one route, then another, the AS_PATHs overlap at the origin AS.

Receive R1. Aggregate 192.0.2.0/24 AS_PATH "64501"

Receive R4. Aggregate 192.0.2.0/24 AS_PATH "[64504 64501]"

If brief aggregation is in use, the AS_PATH is truncated to "".

The resulting AS_PATH is thus not stable and depends on the presence of specific routes.

B.3. Scenario 3: First one route, then another, the AS_PATHs overlap at the neighbor AS

Receive R3. Aggregate 192.0.2.0/24 AS_PATH "64504 64501".

Receive R4. Aggregate 192.0.2.0/24 AS_PATH "64504 [64501 64502]"

If brief aggregation is in use, the AS_PATH is truncated to "64504".

The resulting AS_PATH is thus not stable and depends on the presence of specific routes.

B.4. Achieving Consistent Origin-AS During Aggregation

In the three scenarios above, the aggregating AS 64500 is using traditional brief aggregation. This results in inconsistent origin ASes as the contributing routes are learned.

The trivial solution to addressing the issue is to simply discard all of the ASes for the contributing routes. In simple BGP aggregation topologies, this is likely the correct thing to do. The AS originating the aggregate, 192.0.2.0/24 in this example, is likely the resource holder for the route in question. In such a case, simply originating the route to its BGP upstream neighbors in the Internet with its own AS, 64500, means that a consistent Route Origin Authorization (ROA) could be registered in the RPKI for this prefix. This satisfies the need for a consistent origin AS.

If the contributing ASes are themselves multihomed to the Internet outside of their connections to AS 64500, then additional ROAs would need to be created for each of the more specific prefixes.

In more complex proxy aggregation scenarios, there may be a desire to permit some stable (i.e., common) portion of the contributing AS_PATHs to be kept in the aggregate route. Consider the case for Scenario 3, where the neighbor AS is the same for both R3 and R4 - AS 64504. In such a case, an implementation may permit the aggregate's brief AS_PATH to be "64504", and a ROA would be created for the aggregate prefix with 64504 as the origin AS.

Authors' Addresses

Warren Kumari
Google, Inc.
1600 Amphitheatre Parkway
Mountain View, CA 94043
United States of America

Phone: +1 571 748 4373
Email: warren@kumari.net

Kotikalapudi Sriram
USA NIST
100 Bureau Drive
Gaithersburg, MD 20899
United States of America

Phone: +1 301 975 3973
Email: ksriram@nist.gov

Lilia Hannachi
USA NIST

100 Bureau Drive
Gaithersburg, MD 20899
United States of America

Phone: +1 301 975 3259
Email: lilia.hannachi@nist.gov

Jeffrey Haas
Juniper Networks, Inc.
1133 Innovation Way
Sunnyvale, CA 94089
United States of America

Email: jhaas@juniper.net