Network Working Group
Internet-Draft
Intended status: Best Current Practice
Expires: 1 February 2026

J. Snijders

T. Fiebig
MPI-INF

M. A. Stucchi
Glevia GmbH

31 July 2025

# Guidance to Avoid Carrying RPKI Validation States in Transitive BGP Path Attributes
## draft-ietf-sidrops-avoid-rpki-state-in-bgp-02

## Abstract

This document provides guidance to avoid carrying Resource Public Key Infrastructure (RPKI) derived Validation States in Transitive Border Gateway Protocol (BGP) Path Attributes. Annotating routes with transitive attributes signaling Validation State may cause needless flooding of BGP UPDATE messages through the global Internet routing system, for example when Route Origin Authorizations (ROAs) are issued, or are revoked, or when RPKI-To-Router sessions are terminated.

Operators SHOULD ensure Validation States are not signaled in transitive BGP Path Attributes. Specifically, Operators SHOULD NOT group BGP routes by their Prefix Origin Validation state into BGP Communities.

## 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 1 February 2026.

Table of Contents

## 1. Introduction

The Resource Public Key Infrastructure (RPKI) [RFC6480] allows for validating received BGP routes. By means of this validation process, routes attain a Route Origin Validation (ROV) state. In the past, some operators and vendors suggested to use BGP Communities [RFC1997] and [RFC8092] to annotate received routes with the local Validation State. Some claim that the practice of signaling Validation States could be useful, for example to IBGP speakers, in order to avoid each IBGP speaker having to perform their own route origin validation.

However, annotating a route with a transitive attribute (based on the Validation State) means that BGP update messages have to be sent to every neighbor when the Validation State changes. This means that when for example Route Origin Authorizations [RFC9582] are issued, or are revoked, or RPKI-To-Router [RFC8210] sessions are terminated, new BGP UPDATE messages will have to be sent for all routes that were previously annotated with a BGP Community associated with a different Validation State. Furthermore, given that BGP Communities are a transitive attribute, such a BGP UPDATE might end up propagating through large portions of the Default-Free Zone (DFZ).

Hence, this document provides guidance to avoid carrying RPKI-derived Validation States in Transitive Border Gateway Protocol (BGP) Path Attributes Section 5 of [RFC4271]. Specifically, operators SHOULD NOT group BGP routes by their Prefix Origin Validation state [RFC6811] into BGP Communities [RFC1997] [RFC8092]. If local technical requirements or the implementation used by an operator necessitates the use of transitive attributes to signal RPKI Validation State, the operator SHOULD ensure that these attributes are removed in NLRI announced to EBGP neighbors. Avoiding the use of BGP Communities to signal RPKI Validation States prevents BGP UPDATE messages from being needlessly flooded into the global Internet routing system.

### 1.1. 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.

## 2. Scope

This document discusses signaling locally significant RPKI Validation States to external BGP neighbors through transitive BGP attributes. This includes operator-specific BGP Communities to signal Validation States, as well as any current or future standardized well-known BGP Communities denoting Validation State (for example as specified for Extended BGP Communities in [RFC8097]).

The guidance in this document applies to all current and future transitive BGP attributes that may be implicitly or explicitly used to signal Validation State to neighbors. Similarly, this guidance also applies to non-ROA validation mechanics based on RPKI, e.g., ASPA [I-D.ietf-sidrops-aspa-profile], BGPSec [RFC8205], and any other future validation mechanic built upon the RPKI.

The document acknowleges that specific operational requirements, such as a BGP implementation used by an operator still being dependent on annotating RPKI Validation State using BGP attributes, may necessitate the use of BGP path attributes to signal RPKI Validation State. If this is the case, the dependent operator SHOULD ensure that these attributes are removed before announcing NLRI to EBGP neighbors.

## 3. Risks of Signaling Validation State With Transitive Attributes

This section outlines the risks of signaling RPKI Validation State using BGP Communities. While the current description is specific to BGP communities, the observations hold similar for all transitive attributes that may be added to BGP routes. Furthermore, this document contains data on the measured current impact of BGP Communities used to signal RPKI Validation States.

### 3.1. Triggers for Large-Scale Validation Changes

This section describes examples on how a large amount of RPKI ROV changes may occur in a short time, leading to generation of a large amounts of BGP Updates.

#### 3.1.1. ROA Issuance

Large-Scale ROA issuance should be a comparatively rare event for individual networks. However, several cases exist where issuance by individual operators or (malicious) coordinated issuance of ROAs by multiple operators may lead to a high route churn, triggering a continuous flow of BGP Update messages caused by operators using transitive BGP attributes to signal RPKI Validation State.

Specifically:

* When one large operator newly starts issuing ROAs for their netblocks, possibly by issuing one ROA with a long maxLength covering a large number of prefixes. This may also occur when incorrectly migrating to minimally covering ROAs [RFC9319], i.e., when the previous ROA is first revoked (see Section 3.1.2) and the new ROAs are only issued after this revocation has been propagated, e.g., due to an operational error or a bug in the issuance pipeline used by the operator.

* When multiple smaller operators coordinate to issue new ROAs at the same time.

* When a CA has been unavailable or unable to publish for some time, but then publishes all updates at once, or - as unlikely as it is - if a key-rollover encounters issues.

#### 3.1.2. ROA Revocation

Large-Scale ROA revocation should be a comparatively rare event for individual networks. However, several cases exist where revocations by individual operators or (malicious) coordinated revocation of ROAs by multiple operators may lead to a high route churn triggering a continuous flow of BGP Update messages caused by operators using transitive BGP attributes to signal RPKI Validation State.

Specifically:

* When one large operator revokes all ROAs for their netblocks at once, for example, when migrating to minimally covering ROAs [RFC9319], or when revoking one ROA with a long maxLength covering a large number of prefixes.

* When multiple smaller operators coordinate to revoke ROAs at the same time.

* When a CA becomes unavailable or unable to publish for some time, e.g., due to the CA expiring ([CA-Outage1], [CA-Outage2], [CA-Outage3], [CA-Outage4]).

#### 3.1.3. Loss of Authoritative Validation Information

Similar to the issuance/revocation of ROAs, the validation pipeline of a relaying party may encounter issues. Issues may occur on the router side or on the validator side, with network connectivity issues having specific impact on either of the two.

While, in general, implementations should not have bugs, operators should not make mistakes, and the network should be reliable, this is usually not the case in practice. Instead, the worst-case of sudden and unexpected, yet unintentional, loss of Validation State is an event that, however unlikely in a specific system, may and will happen. Hence, systems should be resilient in case of unexpected issues, and should not further amplify issues by creating a BGP UPDATE storm.

Below, we provide examples of events for both categories that may lead to the Validation State of routes in one or multiple routers of an operator changing from Valid to NotFound. This list serves illustrative purposes and does not claim completeness.

#### 3.1.3.1. Validator Issues

The following events may impact a validator’s ability to provide validation information to routers:

* The RPKI-To-Router (RTR) service may have to temporarily be taken offline by the relying party operator for maintenance. While operators should, in general, take care to provision sufficient redundancy, critical vulnerabilities may necessitate the immediate simultaneous shutdown of all RTR instances.

* A validator may crash due to bugs when ingesting unexpected data from the RPKI, or run into performance issues due to insufficient available memory or limited I/O performance on the host. In the worst case, especially memory issues, can lead to a flapping validator, e.g., when the system runs out of memory after a few minutes of communicating Validation State to routers.

* Validation state may seemingly lapse due to issues with time synchronization if, e.g., the clock of the validator diverts significantly, starting to consider CA’s certificates invalid.

* The validator may lose its network connectivity in general, or to specific CAs. While, in general, the validator should be able to serve from cache, an operator may have to shutdown the validator in such a case, to prevent dropping prefixes as invalid due to stale data.

#### 3.1.3.2. Router Ingestion Issues

* The RTR client, especially when implemented as a dedicated daemon, may fail to start, or terminate when receiving unexpected data. Especially when this leads to a flapping client, e.g., due to a bug in the handling of incremental updates leading to a crash, while the initial retrieval is successful, this will lead to flapping between routes being Valid and NotFound.

* A misconfiguration may impact a router’s ability to communicate with the RTR service. For example, the RTR client may lose its credentials or may not receive updated credentials in time when these are changed, or the address of the RTR service changes and is not updated on the router in time.

* An RTR client may lose network connectivity to the RTR service. While, in general, caches should prevent this from having immediate impact, an RTR clients behavior in case of a flapping network connection with frequent interruptions may lead to unexpected behavior and cache invalidation. Similarly, after cache expirery, routes will change from Valid to NotFound.

* As an extension of the previous point, multiple operators might be using one central RTR service hosted by an external party, or depend on a similar validator, which becomes unavailable, e.g., due to maintenance or an outage. If local instances are not able to handle loss of this external service without changing Validation State, i.e., do not serve from cache or the outage extends beyond cache expirery, routes will change their Validation State from Valid to NotFound Naturally, the negative impact in such a case is significantly larger in comparison to each operator running their own validator.

#### 3.1.4. Outage Scenario Summary

The above non-exhaustive listing suggests that issues in general operations, CA operations, and RPKI cache implementations simply are unavoidable. Hence, Operators MUST plan and design accordingly.

### 3.2. Scaling issues

For each change in Validation State of a route, an Autonomous System in which the local routing policy sets a BGP Community based on the ROV-Valid Validation State, routers would need to send BGP UPDATE messages for more than half the global Internet routing table if the Validation State changes to ROV-NotFound. The same, reversed case, would be true for every new ROA created by the address space holders, whereas a new BGP update would be generated, as the Validation State would change to ROV-Valid.

Furthermore, adding additional attributes to routes increases their size and memory consumption in the RIB of BGP routers. Given the continuous growth of the global routing table, in general, operators should be conservative regarding the additional information they add to routes.

### 3.3. Flooding and Cascading of BGP UPDATES

The aforementioned scaling issues are not confined to singular UPDATE events. Instead, changes in Validation State may lead to floods and/or cascades of BGP UPDATES throughout the Internet.

#### 3.3.1. Flooding of BGP UPDATES

Flooding events are caused by an individual operator losing Validation State. If that operator annotates Validation State using BGP communities, the operator will send updates for all routes that changed from Valid to NotFound to its downstreams, as well as updates for routes received from downstreams to its upstreams.

Following an RPKI service affecting outage (Section 3.1), given that more than half the global Internet routing table with close to 1,000,000 prefixes [CIDR_Report] nowadays is covered by RPKI ROAs [NIST], such convergence events represent a significant burden. See [How-to-break] for an elaboration on this phenomenon.

#### 3.3.2. Cascading of BGP UPDATES

For events that are not specific to one operator, e.g., a malicious widthdrawel of a ROA, loss of a major CA, or an unexpected downtime of a major centralized RTR service, events can also cascade for ASes annotating Validation State using BGP communities. Given that routers' view of the RPKI with RTR are only loosely consistent, update messages may cascade, i.e., one event affecting Validation State may actually trigger multiple subsequent BGP UPDATE floods.

Assume, for example, that AS65536 is a downstream of AS65537 (both annotating Validation State with BGP Communities and using a 300 second RTR cycle), and a centralized RTR service fails. In the example, AS65536 has their routers updated from that cache a second before the service went down, while AS65537 was due for a refresh a second thereafter.

This means that a second after the RTR service went down, AS65537 will trigger a BGP UPDATE flood down its cone. AS65536 will ingest and propagate these BGP UPDATES down its own cone as well.

When, rughly 300 seconds later, AS65536 fails to retrieve Validation State as well, the community set by AS65536 will again change for ROA covered routes, and it will again trigger a BGP UPDATE flood and propagate this down its cone.

Even if either or both of AS65536 and AS65537 use a cache after RTR expirery, the underlying issue would not change, assuming the RTR service downtime spans beyond the cache TTL. Assuming a 30 minute cache TTL, both ASes using a cache would only move the cascading event 30 minutes later. If only one of the two uses a cache, the two flood events get moved further apart. However, the overall issue of two independent floods due to one event remains.

### 3.4. Observed data

In February 2024, a data-gathering initiative [Side-Effect] reported that between 8% and 10% of BGP updates seen on the Routing Information Service - RIS, contained well-known communities from large ISPs signaling either ROV-NotFound or ROV-Valid BGP Validation states. The study also demonstrated that the creation or removal of a ROA object triggered a chain of updates in a period of circa 1 hour following the change.

Such a high percentage of unneeded BGP updates constitutes a considerable level of noise, impacting the capacity of the global routing system while generating load on router CPUs and occupying more RAM than necessary. Keeping this information inside the realms of the single autonomous system would help reduce the burden on the rest of the global routing platform, reducing workload and noise.

### 3.5. Lacking Value of Signaling Validation State

RTR has been developed to communicate validation information to routers. BGP Attributes are not signed, and provide no assurance against third parties adding them, apart from BGP communities--ideally--being filtered at a networks edge. So, even in IBGP scenarios, their benefit in comparison to using RTR on all BGP speakers is limited.

For EBGP, given they are not signed, they provide even less information to other parties except introspection into an ASes internal validation mechanics. Crucially, they provide no actionable information for BGP neighbors. If an AS validates and enforces based on RPKI, Invalid routes should never be imported and, hence, never be send to neighbors. Hence, the argument that adding Validation State to communities enables, e.g., downstreams to filter RPKI Invalid routes is mute, as the only routes a downstream should see are NotFound and Valid. Furthermore, in any case, the operators SHOULD

run their own validation infrastructure and not rely on centralized services or attributes communicated by their neighbors. Everything else circumvents the purpose of RPKI.

## 4. Advantages of Dissociating Validation States and BGP Path Attributes

As outlined in Section 3, signaling Validation State with transitive attributes carries significant risks for the stability of the global routing ecosystem. Not signaling Validation State, hence, has tangible benefits, specifically:

* Reduction of memory consumption on customer/peer facing PE routers (less BGP communities == less memory pressure).

* No effect on the age of a BGP route when a ROA or ASPA [I-D.ietf-sidrops-aspa-profile] is issued or revoked.

* Avoids having to resend, e.g., more than 500,000 BGP routes towards BGP neighbors (for the own cone to peers and upstreams, for the full table towards customers) if the RPKI cache crashes and RTR sessions are terminated, or if flaps in validation are caused by other events.

Hence, operators SHOULD NOT signal RPKI Validation State using transitive BGP attributes.

## 5. Security Considerations

BGP is not guranteed to converge, and the view on the RPKI within an individual administrative domain is only loosely consistent. External validation state anotated in a received NLRI may either depend on a different view on the RPKI than the one in the local administrative domain, or the NLRI may be several hours old itself. Hence, the Validation State of a received announcement can only have local scope.

Additionally, the use of transitive attributes to signal RPKI Validation State may enable attackers to cause notable route churn. This can be accomplished by an attacker issuing and withdrawing, e.g., ROAs for their prefixes, or by the attacker continuously altering transitive attributes used to signal RPKI Validation State for NLRI they readvertise. The latter is possible as NLRI carry no information allowing an ingesting party to validate the integrity of transitive BGP attributes.

DFZ routers may not be equipped to handle route churn in all directions at global scale, especially if said route churn cascades, persists, or repeats periodically. To prevent global route churn,

operators SHOULD NOT signal RPKI Validation State to EBGP neighbors through transitive BGP path attributes. If an operator is dependent on signaling RPKI Validation State among BGP speakers within their AS, they SHOULD ensure that these attributes are removed before announcing NLRI to EBGP neighbors.

Given their potential negative impact, operators SHOULD remove attributes used to signal RPKI Validation State when importing NLRI with an idempotent operation until the corresponding neighbor follows guidance in this document as well.

## 6. IANA Considerations

None.

## 7. Acknowledgements

The authors would like to thank Aaron Groom and Wouter Prins for their helpful review of 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>.

[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>.

### 8.2. Informative References

[CA-Outage1] ARIN, "RPKI Service Notice Update", August 2020, <https://www.arin.net/announcements/20200813/>.

[CA-Outage2] RIPE NCC, "Issue affecting rsync RPKI repository fetching", April 2021, <https://www.ripe.net/ripe/mail/archives/routing-wg/2021-April/004314.html>.

[CA-Outage3]
Snijders, J., "problemas con el TA de RPKI de LACNIC", April 2023, <https://mail.lacnic.net/pipermail/lacnog/2023-April/009471.html>.

[CA-Outage4]
Snijders, J., "AFRINIC RPKI VRP graph for November 2023 - heavy fluctuations affecting 2 members", November 2023, <https://lists.afrinic.net/pipermail/dbwg/2023-November/000493.html>.

[CIDR_Report]
Huston, G., "CIDR REPORT", January 2024, <https://www.cidr-report.org/as2.0/>.

[How-to-break]
Snijders, J., "How to break the Internet: a talk about outages that never happened", CERN Academic Training Lecture Regular Programme; 2021-2022, March 2022, <https://cds.cern.ch/record/2805326>.

[I-D.ietf-sidrops-aspa-profile]
Azimov, A., Uskov, E., Bush, R., Snijders, J., Housley, R., and B. Maddison, "A Profile for Autonomous System Provider Authorization", Work in Progress, Internet-Draft, draft-ietf-sidrops-aspa-profile-19, 6 January 2025, <https://datatracker.ietf.org/doc/html/draft-ietf-sidrops-aspa-profile-19>.

[NIST]
NIST, "NIST RPKI Monitor", January 2024, <https://rpki-monitor.antd.nist.gov/>.

[RFC1997]
Chandra, R., Traina, P., and T. Li, "BGP Communities Attribute", RFC 1997, DOI 10.17487/RFC1997, August 1996, <https://www.rfc-editor.org/info/rfc1997>.

[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>.

[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>.

[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>.

[RFC8092] Heitz, J., Ed., Snijders, J., Ed., Patel, K., Bagdonas, I., and N. Hilliard, "BGP Large Communities Attribute", RFC 8092, DOI 10.17487/RFC8092, February 2017, <https://www.rfc-editor.org/info/rfc8092>.

[RFC8097] Mohapatra, P., Patel, K., Scudder, J., Ward, D., and R. Bush, "BGP Prefix Origin Validation State Extended Community", RFC 8097, DOI 10.17487/RFC8097, March 2017, <https://www.rfc-editor.org/info/rfc8097>.

[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>.

[RFC8210] Bush, R. and R. Austein, "The Resource Public Key Infrastructure (RPKI) to Router Protocol, Version 1", RFC 8210, DOI 10.17487/RFC8210, September 2017, <https://www.rfc-editor.org/info/rfc8210>.

[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>.

[RFC9582] Snijders, J., Maddison, B., Lepinski, M., Kong, D., and S. Kent, "A Profile for Route Origin Authorizations (ROAs)", RFC 9582, DOI 10.17487/RFC9582, May 2024, <https://www.rfc-editor.org/info/rfc9582>.

[Side-Effect] Stucchi, M., "A BGP Side Effect of RPKI", February 2024, <https://labs.ripe.net/author/stucchimax/a-bgp-side-effect-of-rpki/>.

## Authors' Addresses

Job Snijders
Amsterdam
Netherlands
Email: job@sobornost.net

Tobias Fiebig
Max-Planck-Institut fuer Informatik
Campus E14
66123 Saarbruecken
Germany
Phone: +49 681 9325 3527
Email: tfiebig@mpi-inf.mpg.de

Massimiliano Stucchi
Glevia GmbH
CH- Bruettisellen
Switzerland
Email: stucchi@glevia.ch