sidrops
Internet-Draft
Intended status: Best Current Practice
Expires: 31 August 2025

J. Zhang
Zhongguancun Laboratory
M. Xu
Y. Wang
Tsinghua University

27 February 2025

Enhancing Route Origin Validation by Aggregating Validated ROA Payloads
draft-zhang-sidrops-vrp-aggregation-02

Abstract

Resource Public Key Infrastructure (RPKI) enables address space holders to issue Route Origin Authorization (ROA) objects to authorize one or more Autonomous Systems (ASes) to originate BGP routes for specific IP address prefixes. Individual BGP speakers can utilize Validated ROA Payloads (VRPs) to validate the legitimacy of BGP announcements. This document highlights potential validation errors, and recommends extension of VRPs from reasonable aggregation to enhance Route Origin Validation (ROV) and prevent validation errors that may occur due to traffic engineering or route aggregation policies.

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 31 August 2025.

Copyright Notice

Table of Contents

1. Introduction

In Resource Public Key Infrastructure (RPKI), an address space holder issues a digitally signed object called Route Origin Authorization (ROA) to authorize a specific Autonomous System (AS) to announce BGP routes for one or more IP prefixes within the corresponding address space. The BGP speaker loads validated RPKI ROA objects from the Relying Party (RP) cache into its local storage. The loaded objects are formatted with (prefix, originating AS number, maximum length). These locally stored objects are referred to as "Validated ROA Payload" (VRP), as defined in [RFC6811]. VRPs will be used to validate the origination AS of BGP routes[RFC6483] .

However, due to factors such as traffic engineering or route aggregation, the prefixes announced by a BGP route may not be consistent with the prefixes registered in the ROA. Typically, address space holders register specific prefixes in the ROA, while the BGP route announcements may involve aggregated prefixes.

This document proposes to extend the original VRPs with new VRPs, which are generated based on aggregation of contiguous prefixes authorized to the same origin AS in original VRPs. VRP aggregation could improve the accuracy of route announcement validation, prevent valid announcements from being erroneously validated as "Invalid" or "Unknown," and avoid unnecessary traffic discarding. Ultimately, this approach aims to improve the practicality and effectiveness of RPKI deployment.

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. Problem Statement and Root Cause Analysis

RPKI is a cryptographic system that enables validation of the origin of route announcements and helps prevent route origin hijacking and other security threats. Typically, if an address space holder does not register a specific prefix in RPKI, it implies that the holder does not authorize an AS to advertise that prefix in routing announcements.

However, there are situations where a prefix included in a route announcement may be subject to aggregation or deaggregation due to factors such as traffic engineering or route optimization. As a result, the prefix being advertised might differ from the registered specific prefix in RPKI.

In such scenarios, when a route validation process relies solely on RPKI ROAs, it may inaccurately validate the route announcement as "Invalid" or "Unknown". This can happen when the aggregated parent prefix is announced, yet only the pre-aggregation sub-prefixes are registered in the RPKI. Complex routing policies or inaccurate registrations can both lead to similar situations.

In this section, we outline a series of typical scenarios that may lead to the validation outcomes being erroneously labeled as "Invalid" or "Unknown". We will explore three potential causes for

such inaccuracies associated with aggregated but unregistered prefixes, specifically: route aggregation, traffic engineering (including path redundancy, load balancing, etc.), and issues arising from inaccurate registration.

2.1. Route Aggregation

By merging a series of contiguous IP address prefixes into a single less-specific prefix, routing information can be simplified and routing efficiency improved. However, if the AS that performs route aggregation has only been authorized to origin multiple sub-prefixes, the resulting aggregated BGP route announcement, will not be validated as "Valid".

* Example. As shown in Figure 1, the BGP route announcement for the prefix 76.191.76.0/22 from AS62915 is erroneously validated as "Invalid". AS62915 has announced a prefix 76.191.76.0/22 in global BGP routing table, but it has been authorized to originate a set of contiguous sub-prefixes across three ROAs (76.191.74.0/23, AS62915, maxlength=24), (76.191.76.0/23, AS62915, maxlength=24), and (76.191.78.0/23, AS62915, maxlength=24). These registered sub-prefixes can be aggregated into the parent prefix 76.191.76.0/22. However, these ROAs alone cannot validate the route 76.191.76.0/22 originating from AS62915. Worse still, the upstream provider of AS62915, AS11404, has been authorized to originate a more extensive range of prefixes (76.191.64.0/18, AS11404, maxlength=24) in RPKI ROA, which includes the prefix 76.191.76.0/22. Consequently, the legitimate route announcement for the prefix 76.191.76.0/22 originating from AS62915 is mistakenly validated as "Invalid". This incorrect validation results in the route announcement being discarded, leading to traffic intended for AS62915 not being routed correctly.

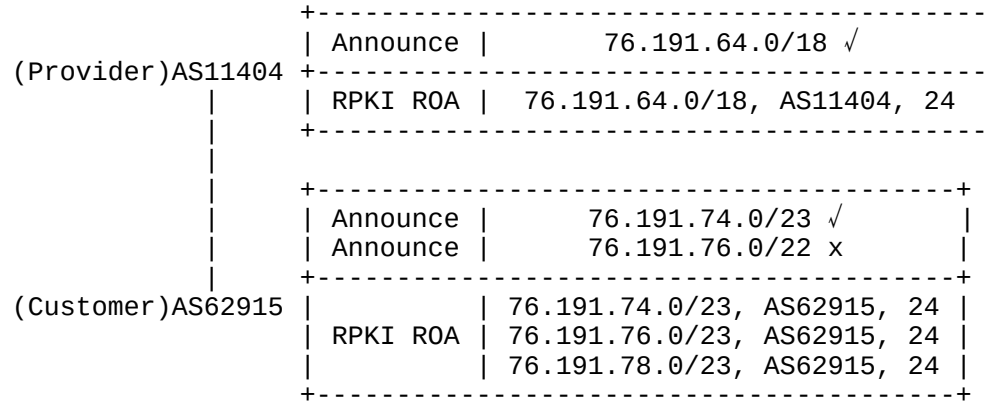


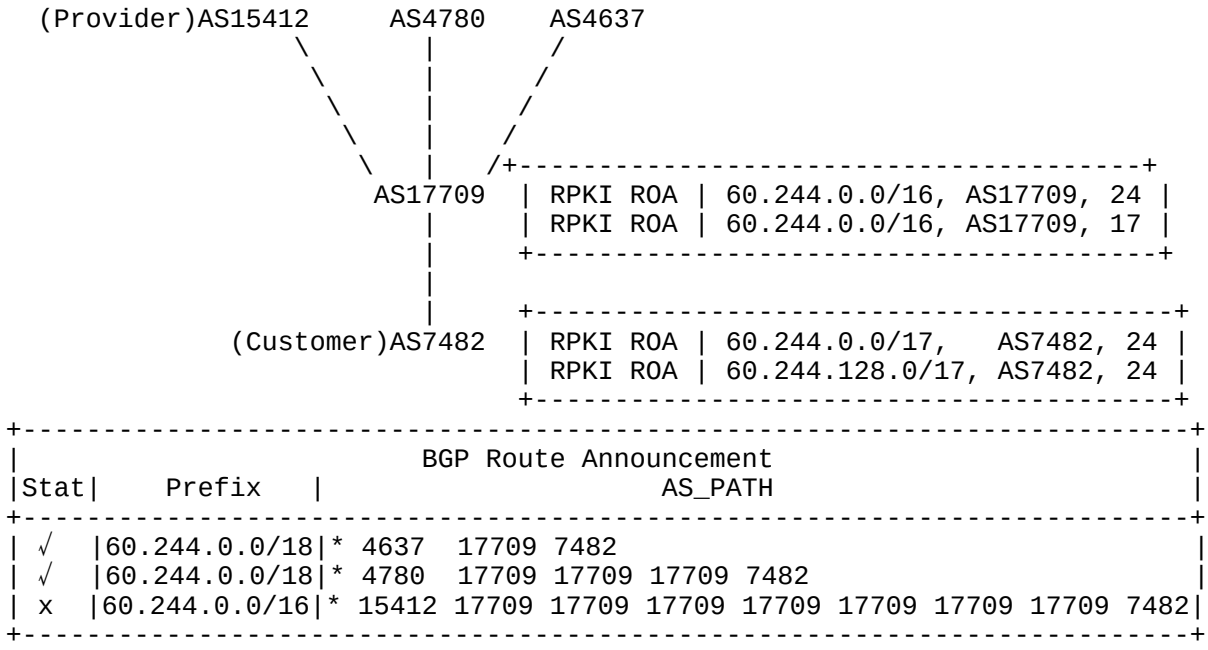
Figure 1: Prefix 76.191.76.0/22 announced by AS62915 is erroneously validated as "Invalid". A '√' symbol following the prefix indicates that it has been validated as "Valid", while a 'x' symbol indicates "Invalid".

2.2. Traffic Engineering

Internet Service Providers (ISPs) might utilize traffic engineering to enhance the network resource utilization, optimize network performance, and ensure Quality of Service (QoS). In this context, operators might announce multiple prefixes with parent-child relationships for the following considerations. If only the sub-prefixes are registered in RPKI, the parent prefix will be unable to be validated.

* Path Redundancy. In the scenario of multi-homing, child prefixes are announced to certain providers, while the parent prefix is announced to others. The parent prefix path can serve as a backup path. Thus, if the child prefix becomes unreachable, traffic can still be routed via the parent prefix. The parent route 60.244.0.0/16 from AS7482 in Figure 2 works as a backup path. If there are failures in the path of the route announcement for the sub-prefix 60.244.0.0/18, the existence of the announcement of 60.244.0.0/16 ensures that traffic can still be routed to AS7482 via providers AS15412 and AS17709, maintaining network stability and reachability. However, despite AS7482 announcing the prefix 60.244.0.0/16 in the global BGP routing table, it has only been authorized to originate a set of contiguous sub-prefixes in two ROAs (60.244.0.0/17, AS7482, maxlength=24) and (60.244.128.0/17, AS7482, maxlength=24). These two registered sub-prefixes can be aggregated into the parent prefix 60.244.0.0/16. However, since there are two ROAs (60.244.0.0/16, AS17709, maxlength=24) and

(60.244.0.0/16, AS17709, maxlength=17) that authorized the prefix 60.244.0.0/16 to be originated from AS17709, the provider of AS7482, this BGP route anouncement is erroneously validated as "Invalid".



BGP Route Announcement		
Stat	Prefix	AS_PATH
√	60.244.0.0/18	* 4637 17709 7482
√	60.244.0.0/18	* 4780 17709 17709 17709 7482
x	60.244.0.0/16	* 15412 17709 17709 17709 17709 17709 17709 17709 7482

Figure 2: Prefix 60.244.0.0/16 announced by AS7482 is erroneously validated as "Invalid". The '*' symbol in AS_PATH indicates other ASes in the path that are irrelevant to the traffic engineering policy. From top to bottom, the relationships are in a provider-customer hierarchy.

* Load Balancing. If an AS is connected to multiple upstream providers, it may announce different parent and child prefixes through different providers to achieve fine-grained traffic distribution, thus accomplishing load balancing. As shown in Figure 3, the parent prefix 93.113.148.0/22 is only announced through AS6762, while its sub-prefix 93.113.150.0/24 is announced through three providers, AS6939, AS2914, and AS6762, which could achieve load balancing. However, AS49367 has not been authorized to originate the parent prefix, this BGP route announcement for 93.113.148.0/22 is validated as "Unknown".

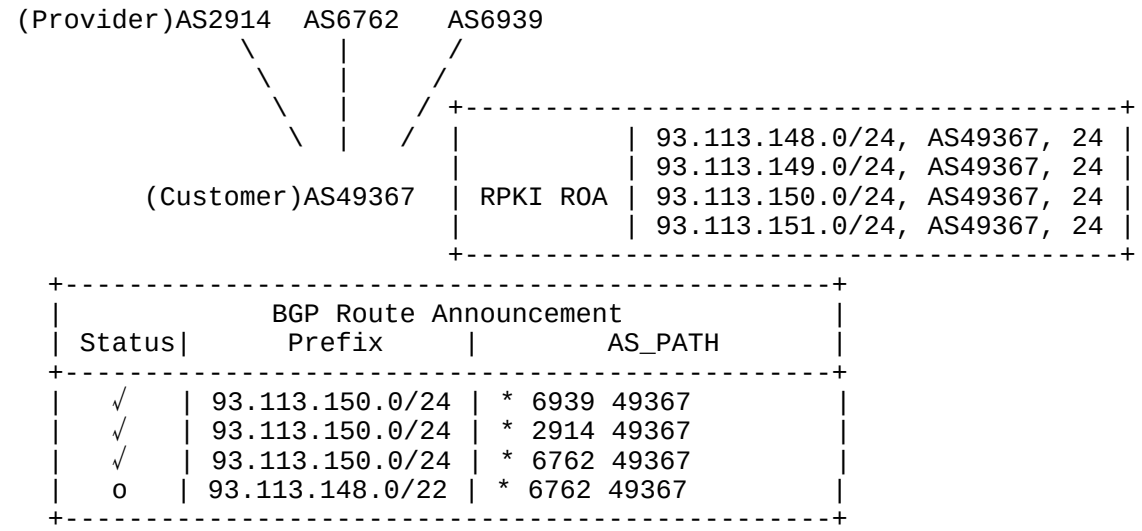


Figure 3: Prefix 93.113.148.0/22 announced by AS49367 is erroneously validated as "Unknown". A ’o’ symbol following the prefix indicates that it has been validated as "Unknown", while a ’√’ symbol indicates "Valid".

* Traffic shaping, performance optimization, etc. For reasons such as traffic shaping and performance optimization, an AS may announce multiple prefixes that have a parent-child relationship. In this case, if the AS has only been authorized to originate the sub-prefixes it owns, this would result in the parent prefix not being validated as "Valid".

2.3. Incomplete or inaccurate registration

RPKI is still in the process of incremental deployment; therefore, some prefixes have not yet been registered. Additionally, some ISPs do not maintain consistency between the ROAs registered in RPKI and the prefixes they announce. They may also fail to register the aggregated parent prefix in a timely manner after implementing an aggregation policy, leading to validation issues.

3. Algorithm of VRP Aggregation

To address the problem described above, this document suggests extending VRPs by aggregating contiguous prefixes from the same AS into new VRPs. This section introduces a preliminary algorithm for VRP aggregation and details the implementation process, ensuring the correctness of the newly aggregated VRP.

3.1. Algorithm Rationality

This document suggests extending VRPs by aggregating those that contain contiguous prefixes authorized to the same AS. In this context, "contiguous" refers to IP address ranges that are sequentially contiguous in binary representation. This implies that the IP address spaces represented by the prefixes do not overlap with each other and there are no gaps between them.

The content of the ROA identifies that an IP address block holder has authorized an AS to announce routes to one or more prefixes within the address block. If an AS is authorized to several multiple contiguous IP address prefixes, it signifies that the AS can facilitate route reachability for the IP address blocks corresponding to these prefixes. Therefore, if this AS announces a parent prefix that aggregates these contiguous sub-prefixes, the routing to this parent prefix should also be considered accessible and recognized as a "Valid" route announcement.

The address holder is often not aware of network routing policies when issuing ROAs. The current RPKI architecture does not provide a mechanism to accommodate these dynamic changes in the registered ROAs. Therefore, this document advocates for the development of specialized extending the aggregated VRP aggregation to address this issue and achieve the aforementioned goal.

3.2. Aggregatable ROA Payload

The fundamental principle of VRP Aggregation is to ensure the correctness of the newly generated prefixes after aggregation. In line with this principle, this document puts forward three rules for VRP aggregation.

1. Only contiguous prefixes from the same AS can be aggregated.

2. Only contiguous prefixes with the same maxLength can be aggregated, taking into account the address holder’s specific consideration of "maxLength" during ROA registration. The resulting VRP will feature the aggregated parent prefix and will inherit the "AS number" and "maxLength" values from the original VRPs.

3. All aggregatable prefixes will be aggregated into a single, largest parent prefix. Different sets of child prefixes may result in multiple potential parent prefixes, but only the largest one will be chosen. For example, the four registered sub-prefixes in Figure 3 could be aggregated into either 93.113.148.0/22 or into two separate prefixes: 93.113.148.0/23

and 93.113.150.0/23. However, only 93.113.148.0/22 will be retained, and the aggregated VRP will be represented as (93.113.148.0/22, AS49367, maxlength=24). This approach ensures that if BGP route advertises any other parent prefixes that could potentially be aggregated into(e.g., 93.113.148.0/23), they can still be validated as "Valid" by referencing the largest aggregated ROA (93.113.148.0/22, AS49367, maxlength=24).

4. Implementation of VRP Aggregation

This document provides two implementation options for VRP aggregation. Operators can make choices based on their specific needs and circumstances.

4.1. Router Extensions

The first implementation option is a data plane solution, which recommends feature extensions for routers. An implementation of the VRP Aggregation on routers is outlined as follows.

1. Once synchronizing all the VRPs from RP, the router aggregates the current VRPs and stores the new, aggregated prefixes and AS numbers in the VRP format as specified in [RFC6811]. The original and aggregated VRPs are stored separately.

2. When the router receives a BGP update, it performs BGP Route Origin Validation (ROV) [RFC6811] with the original, unaggregated VRPs.

3. If the validity state of the received BGP route is determined to be "Valid", go to Step 4. If the validity state of the received BGP route is "Invalid" or "Unknown", the router then performs ROV with the newly aggregated VRPs. If the newly determined validity state is "Valid", the router accepts this BGP route. However, if the new validity state remains "Invalid" or "Unknown", the router ignores this outcome, and the validity state of this BGP route remains as it was determined during the initial validation (Step 2).

4. The router applies the validity state of the BGP route to the route selection [RFC6483].

In the context of VRP aggregation, only the validity state validated as "Valid" by the aggregated VRPs is adopted. This is because the aggregated VRPs only provide certain potential BGP route announcements, indicating that these route announcements are reasonable if they occur. However, it can not validate the state of route announcements originating from other ASes.

For example, as shown in Figure 4, the two sub-prefixes authorized to be originated by AS4809 in two ROAs in RPKI could be aggregated in (202.111.192.0/19, AS4809, maxlength=20). However, AS4809 does not announce the parent prefix 202.111.192.0/19; instead, its provider AS4134 does. In this case, the aggregated VRP may cause the route announcement for the parent prefix to be validated as "Invalid" instead of "Unknown". If the router were to reject this route announcement (202.111.192.0/19), it could disrupt the corresponding traffic. Consequently, the validation results from the aggregated VRPs that are validated as "Invalid" are not accepted.

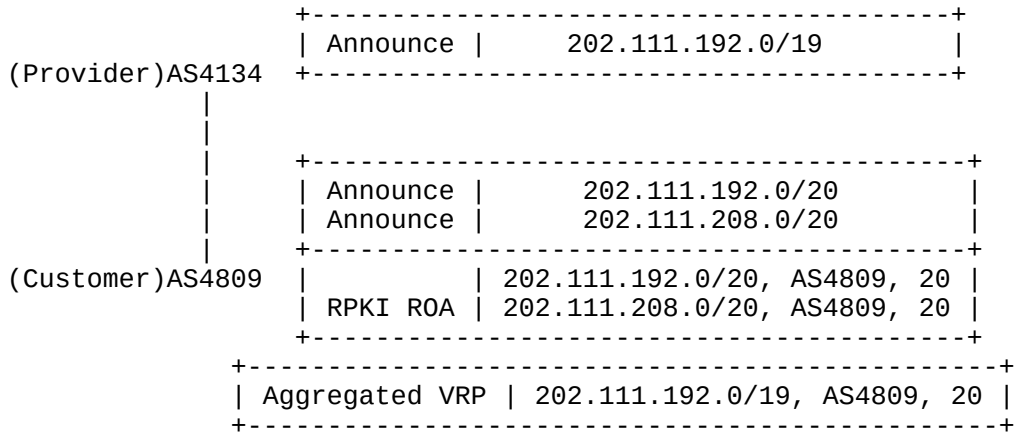


Figure 4: Prefix 202.111.192.0/19 announced by AS4134 is validated as "Invalid" by the aggregated VRP. In such a case, the validity state will not be accepted and it will retain the validity state determined by the original VRPs.

4.2. Relying Party Extensions

The second implementation option is a control plane solution, which requires distributing the routing table to the RP to determine the aggregated VRP used in Route Origin Validation (ROV). An implementation of the VRP Aggregation on RP is outlined as follows.

1. Once synchronizing and validating all the ROAs and generating the original VRPs, RP aggregates the current VRPs and stores the new, aggregated prefixes and AS numbers in the VRP format as specified in [RFC6811]. The original and aggregated VRPs are stored separately.

2. Each VRP/ROA represents an AS authorized to announce one or more prefixes. Compared to the original VRPs, the aggregated VRP implies that ASes can legitimately announce additional prefixes. RP records origin ASes and the prefixes they are newly added to announce legitimately.

3. From the recorded orgin AS and prefix pairs, RP extracts the pairs that appear in the routing table, expands them into the VRP format and adds them to the original VRP. Here, the maxlength is set to the prefix length.

4. The expanded original VRP will be delivered to the router via the RPKI to Router Protocol [I-D.ietf-sidrops-8210bis], enabling BGP ROV [RFC6811] execution.

Similar to the implementation scheme on routers, for the prefixes represented by the aggregated VRP, only the prefixes that match the routes in the routing table will be delivered to the router. Unlike the implementation scheme on routers, in the case of MOAS, the newly added VRP may validate the matched route as "Valid", while simultaneously validating another route announced by a different AS as "Invalid". In this situation, the owner of the address space needs to proactively authorize the prefix to all the ASes to avoid this issue.

4.3. Considerations for VRP Aggregation Implementation

The usage of aggregated VRP is different from the original VRP. The aggregated VRPs serve as a retrospective correction mechanism that supplements some potentially valid route announcements. If there is a route announcement whose prefix and origin AS match the aggregated VRP, it allows that route announcement to be validated as "Valid". However, the presence of an aggregated VRP does not necessarily mean that the AS will announce the aggregated prefix, nor could it validate any other route announcements that do not match as "Invalid". Consequently, one approach is that routers require modifications to employ the ROV process differently for aggregated VRPs than for the original VRPs. Another approach is that RPs determine which aggregated VRPs can be used to perform ROV, based on the routes on the BGP router in some way, as previously described.

In fact, such a situation can be avoided if the address space holders register all of the route announcements they may advertise in RPKI. For instance, AS4134 in Figure 4 could be authorized to originate the prefix 202.111.192.0/19. However, given the current low rate of ROA registration in RPKI, we choose not to adopt the instances where the aggregated VRPs could not validate a route as "Valid", to avoid unnecessary discarding of traffic.

5. Security Considerations

TBD.

6. IANA Considerations

TBD.

7. References

7.1. 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>.

[RFC6483] Huston, G. and G. Michaelson, "Validation of Route Origination Using the Resource Certificate Public Key Infrastructure (PKI) and Route Origin Authorizations (ROAs)", RFC 6483, DOI 10.17487/RFC6483, February 2012, <https://www.rfc-editor.org/info/rfc6483>.

[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>.

[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>.

[I-D.ietf-sidrops-8210bis] Bush, R. and R. Austein, "The Resource Public Key Infrastructure (RPKI) to Router Protocol, Version 2", Work in Progress, Internet-Draft, draft-ietf-sidrops-8210bis-14, 6 August 2024, <https://datatracker.ietf.org/doc/html/draft-ietf-sidrops-8210bis-14>.

7.2. Informative References

[RFC6481] Huston, G., Loomans, R., and G. Michaelson, "A Profile for Resource Certificate Repository Structure", RFC 6481, DOI 10.17487/RFC6481, February 2012, <https://www.rfc-editor.org/info/rfc6481>.

[RFC8897] Ma, D. and S. Kent, "Requirements for Resource Public Key Infrastructure (RPKI) Relying Parties", RFC 8897, DOI 10.17487/RFC8897, September 2020, <https://www.rfc-editor.org/info/rfc8897>.

Acknowledgements

TBD.

Authors' Addresses

Jia Zhang
Zhongguancun Laboratory
Email: zhangj@mail.zgclab.edu.cn

Mingwei Xu
Tsinghua University
Email: xmw@cernet.edu.cn

Yangyang Wang
Tsinghua University
Email: wangyy@cernet.edu.cn