IIJ's Efforts with RPKI

2.1 What is Route Hijacking?

The Internet is formed by the interconnection of organizations (networks) identified by 2-byte or 4-byte AS (Autonomous System) numbers (e.g., IIJ's AS is 2497). The ASes are connected by a routing protocol called BGP (Border Gateway Protocol), and each AS advertises its own IP address to one another in the form of route information. This information propagates around the world and thus provides a mechanism by which packets can arrive at a destination from the other side of the globe.

The IP addresses each AS uses are strictly controlled by the RIRs (Regional Internet Registries; in Asia this is APNIC (Asia-Pacific Network Information Center)) delegated for each region by IANA (Internet Assigned Numbers Authority) and by the NIRs (National Internet Registries; in Japan, this is JPNIC (Japan Network Information Center)) for each country. Each AS receives IP address allocations from these authorities. As long as each AS accurately advertises BGP routes only for the addresses it has been allocated, no problems occur, but what happens if an AS, for whatever reason, ends up advertising IP addresses it has not been allocated? For example, naturally only IIJ should be advertising the route 202.232.0.0/16, which includes 202.232.2.164, the IPv4 address of the IIJ website (www.iij.ad.jp). But if an AS somewhere that is not IIJ were to advertise 202.232.2.0/24, which is part of the above route, packets intended for the IIJ homepage will reach this AS (a principle of routing is that routes with longer lengths take priority). There is not really much of an impact in the case of the IIJ website, but it is easy to imagine what the impacts could be in the case of DNS servers or banking sites.

This phenomenon is generally called route hijacking, and these sorts of issues do actually happen on the Internet on a daily basis. Examples include prominent video site YouTube's service being suspended because an AS that is not Google advertised a certain route, and incidents in which BitCoin is said to have been misappropriated when routes encompassing BitCoin-related site addresses were advertised by a separate AS. So how do problems like this arise? Each AS essentially self-declares the aforementioned BGP route advertisements. Confirming that the routes advertised by the AS of the system you are connected to are legitimate is utterly infeasible as it would require routers to reflect the innumerable IP allocations that are updated daily, so there is no choice but to almost unconditionally accept the advertised routes. So in some sense, the Internet as it currently stands is on quite precarious footing.

2.2 Overview of RPKI

With the Internet having now become an indispensable part of our social infrastructure, leaving the situation unaddressed would expose society as a whole to considerable risk, so RPKI (Resource Public-Key Infrastructure) has been devised to rectify this. The idea of RPKI appeared circa 1998, around the time the Internet finally became widespread in Japan, and it represents amazing foresight on the part of researchers.

In a nutshell, RPKI provides a mechanism for verifying/validating the legitimacy of resources (Internet number resources such as IP addresses and AS numbers) using digital certificates (X.509). As mentioned, IP address allocations are managed by IANA, RIRs, and NIRs, so these operating organizations form a tree structure (to be precise, a tree with five RIRs at the top), and digital certificates guarantee that the resources are correct. Users of the information use these digital certificates to determine that the resources are correct. RPKI itself is a general-purpose mechanism that is also applicable to scenarios beyond BGP routing, but we limit our discussion here to BGP routing.

An AS, having received an IP address allocation, registers the IP addresses for which it intends to advertise BGP routes, along with the maximum prefix length and the origin AS number, with the RPKI system managed by its NIR. The RPKI system issues a digital certificate in response[*1]. This digital certificate is called a ROA (Route Origination Authorization).

Users of these ROAs rely on information called TALs (Trust Anchor Locators), which correspond to the vertices of a tree structure, to traverse the tree up through the NIR and RIR and acquire the ROA, which they then verify and save as validated data (VRP, Validated ROA Payload). It is the role of a cache server to provide this VRP to the router. The information is supplied to the BGP router via a protocol called RPKI-RTR (RPKI to Router Protocol). Based on this information, the BGP router verifies whether route advertisements it receives are correct by matching their content up against the VRP data. Consider, for instance, a VRP with IP address 202.232.0/16, maximum length /17, and ASN 2497. A route advertisement with IP address 202.232.2.0/24 and ASN 64494 would be invalid, and refusing to accept this route can prevent route hijacking. Validating the origin AS on received routes using RPKI information (ROA) like this is called ROV (Route Origin Validation). How the validation results are handled is left up to the operating policies of each AS, but common practice at present is to discard only those routes that are clearly invalid (for reasons explained below).

2.3 Current State of RPKI

As of January 2021, BGP route information for around 930,00 routes (IPv4 830,000, IPv6 100,000) is being exchanged on the Internet, but the number of valid ROAs stands at about 210,000. The ROA count as of October last year was roughly 190,000, so it has increased by 20,000 in four months and is thus right in the middle of its expansion. Figure 1 shows routes that can and routes that cannot be validated using ROA as a proportion of all BGP routes (930,000). Although the number is steadily increasing, over 70% of BGP routes do not yet have a ROA, so it is not possible to validate the originating AS using ROA. Above, I explained that with current ROV, routes are generally only discarded when they are clearly invalid, and this is why. When it is unclear whether a route is proper or not because it cannot be validated, there is no option but to accept it. The hope is that RPKI will continue to spread to that point that all IP addresses can be validated, but that will likely take a decent amount of time.

Of the roughly 71,000 ASes for which BGP routes exist, around 20,000 have ROAs with the AS listed as the origin AS. In the case of the AS with the most routes, ROAs exist for around 4,000 of the roughly 9,600 BGP routes originated by that AS, but this AS has a prefix length / maximum prefix length of /20, and it also has ROAs for this range broken into the prefix lengths /21, /22, /23, and /24. Normally in this case, a single ROA would do with a prefix length of

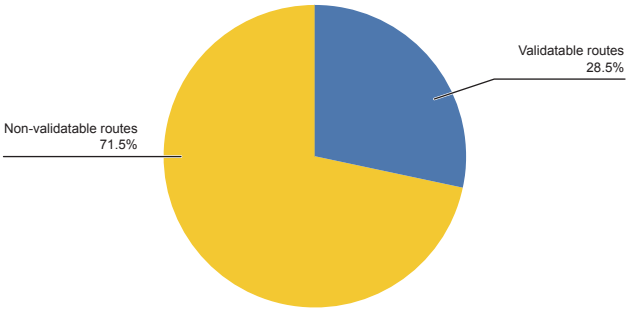

Figure 1: Validatable and Non-validatable Routes

*1 The IP address user can also issue a digital certificate. In this case, the user is the CA (Certification Authority) and will be incorporated into the trust tree as the authority for the IP address assigned by the NRI.

/20 and a maximum prefix length of /24, so it is unclear what is being achieved here, but we can say that creating unnecessary ROAs is not the proper thing to do as it results in the unnecessary consumption of router resources.

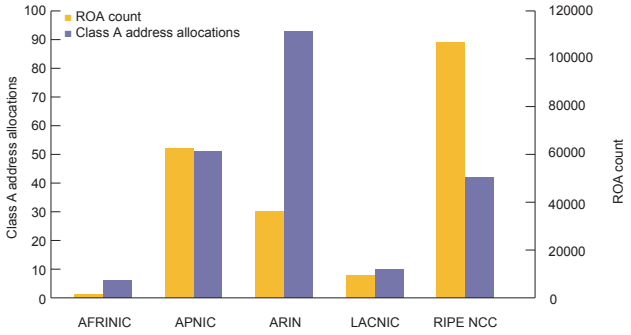

Figure 2: Address Allocations and ROAs by RIR

Next, we look at the state of ROAs by region. Figure 2 shows the number of class A address allocations and ROAs for each RIR*2.

As you can see, APNIC, which oversees the Asian region including Japan, RIPE, which oversees Europe, and LACNIC, which oversees Latin America, create a large number of ROAs relative to the number of allocated addresses. And on a country-by-country basis, it looks like some countries have reached 100%*3. Unfortunately, the adoption rate is not high in Japan, so hopefully we will see greater efforts in this regard ahead.

Let's look at the ROA prefix lengths. Figures 3 and 4 show the distribution of, respectively, IPv4 and IPv6 ROA prefix lengths and maximum prefix lengths. Generally, the usual practice on the Internet is to exchange IPv4 prefix lengths of up to /24 and IPv6 prefix lengths of up to /48, and so routes with prefix lengths longer than those are not exchanged. Yet

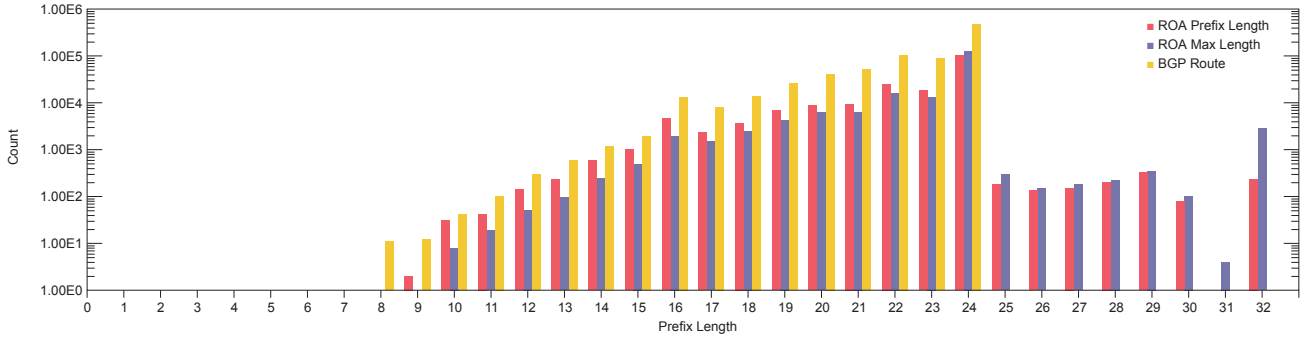

Figure 3: ROA Prefix Lengths and BGP Prefix Lengths (IPv4)

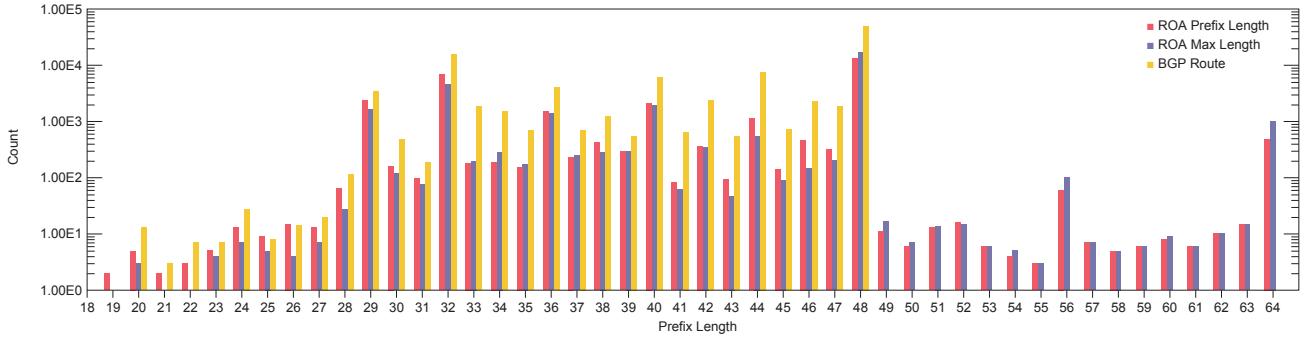

Figure 4: ROA Prefix Lengths and BGP Prefix Lengths (IPv6)

*2 Allocations to each RIR are based on IANA data (https://www.iana.org/numbers). Because international address transfers are made to ensure IPv4 works effectively, differences arise between RIR allocations and the actual region of use, so the figures do not necessarily show the correct region of use.

*3 NLnet Labs, RPKI Tools (https://nlnetlabs.nl/projects/rpki/rpki-analytics/).

there appear to be quite a number of ROAs with long prefix lengths. Further, Figure 5 shows the distribution of the difference between ROA prefix length and maximum length, and while there is no difference in the overwhelming number of cases, there are also substantial differences in quite a few cases. RPKI-based ROV only verifies that the combination of IP address and origin AS is correct; it does not deal with cases in which information, including the origin AS itself, is spoofed. In general, routes become hijacked when an operator originates a route with a prefix that is longer than that of the normal BGP route, and setting a maximum length in a ROA that is longer than the BGP actually being advertised contributes to this risk. So it is best to do everything possible to ensure that advertised BGP routes and ROAs have the same maximum length. But if you do accidentally advertise a prefix longer than the ROA's maximum length, the route will be discarded under ROV, causing a routing failure, so considerable care must be taken.

So far we have looked at the state of ROA. Now let's look at how many invalid routes are detected via ROV using ROAs. As I will explain, IIJ adopted ROV at end-2020, so in principle there are no invalid routes within the IIJ network. We thus use slightly older data and look at the situation around August 2020, before IIJ began using ROV. Figure 6 shows the results of ROV on BGP routes received by IIJ. Along with "valid" and "invalid" results, we also have "not found", which means there was no ROA, so validation is not possible. As indicated, around 3,000 routes, or 0.3% of the total, were invalid as of end-August 2020.

These roughly 3,000 invalid routes are broken down in Figure 7. Around half have the correct origin AS but the wrong prefix length (mismatch length); around 30% have the wrong origin AS (mismatch origin); and the remaining 20% have the wrong origin AS and prefix length (mismatch origin and length). Many of the length mismatches are probably cases in which routes internal to the AS that have long prefix lengths are accidentally advertised externally when they shouldn't be (leaked). The mismatches of both origin AS and length are possibly malicious route hijacking attempts, but there are also likely many cases in which part of an address range allocated to one AS is being advertised by another AS (commonly called hole punching). Where hole

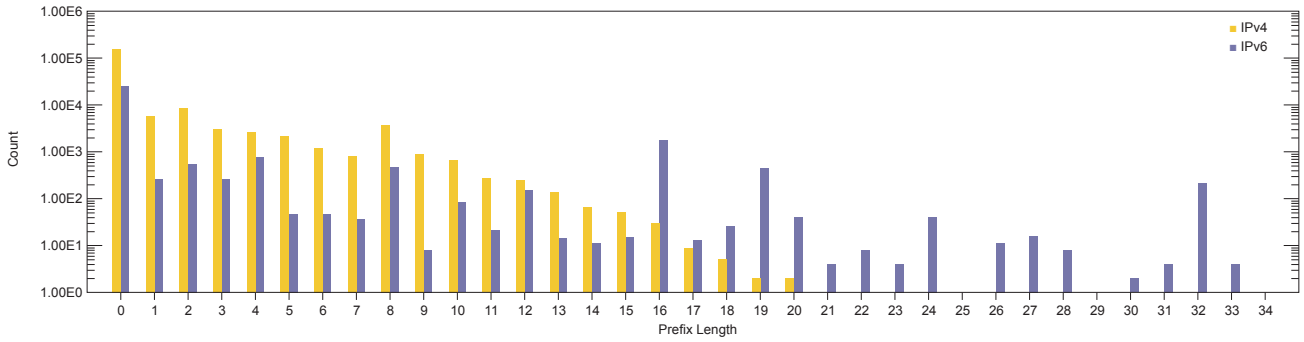

Figure 5: Difference between ROA Prefix Length and Max Length

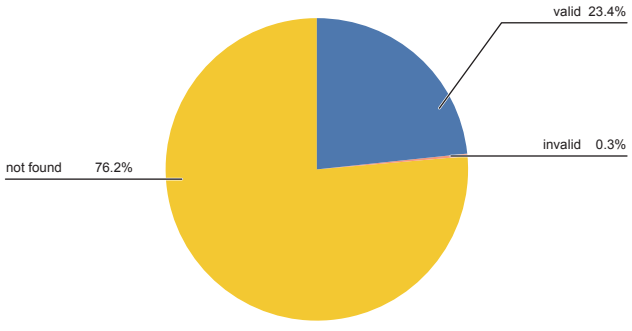

Figure 6: Breakdown of ROV Results

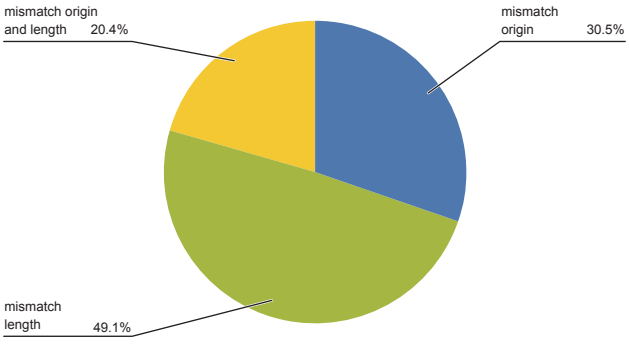

Figure 7: Breakdown of Invalid Routes

punching is occurring, separate ROAs should be created for the different origin ASes and prefix lengths based on the allocated address range and the more specific range, but it is conceivable that someone has neglected to create a ROA for the more specific range. But either way, only the people creating the routes actually know what their real intentions are. From the outside looking in, it's impossible to tell if it's simply an oversight, route hijacking due to a configuration error, or malicious route hijacking. So routes deemed invalid by ROV are uniformly discarded, resulting in a non-zero chance of dropping some routes that should not be dropped. Proper management of ROAs and advertised routes is the responsibility of the AS that receives the IP address range, so there is no fault on the part of ASes that discard routes according to ROV.

This means that roughly 3,000 routes ROV deems invalid are all discarded, but this does not necessarily mean that they all become unreachable. For example, even if 192.0.2.0/25 is discarded, reachability is retained if there is a route for 192.0.2.0/24, which encompasses this. But there are actually many cases in which the origin ASes for, in this example, 192.0.2.0/25 and 192.0.2.0/24 differ, and in such cases, even if 192.0.2.0/24 does exist, it is difficult to objectively determine whether packets reach the proper destination. If cases where the origin ASes differ are permitted, there are alternative routes for around 2,500 of these roughly 3,000 routes, and if only cases in which the origin AS is the same are permitted, there are alternative routes for around 1,500 of them. So taking a strict view, ROV results in reachability being lost for around 500 routes (roughly 0.04% of all BGP routes); and taking a looser view, it results in around 1,500 routes (0.15%) becoming unreachable.

Table 1: Test Routes Retained

Route	RPKI	route views	RIS
93.175.146.0/24	Valid	28 AS	287 AS
93.175.147.0/24	Invalid	13 AS	207 AS
2001:7fb:fd02::/48	Valid	N/A	290 AS
2001:7fb:fd03::/48	Invalid	N/A	205 AS

Naturally, reducing route hijacking itself is the objective of ROV, so discarding invalid routes is the right course of action, but the introduction of ROV does bring with it the possibility of blocking traffic that was previously being routed, even if perhaps improperly. So it would be good to have an idea of what the impact of that would be beforehand.

So how many ASes around the world have adopted ROV? Unlike with ROA, it is difficult to tell for sure from an external perspective whether each AS has adopted AS. Although it depends on self-reporting by ASes, according to the website https://isbgpsafeyet.com/, created to increase awareness of RPKI, around 100 ASes have implemented ROV. More objective information can be gleaned from routes advertised by RIPE NCC, the RIR for Europe, for the purpose of measuring certification technology uptake[*4]. The routes intentionally include both routes designed to be valid and invalid according to ROV, so the degree to which ASes retain these routes can be used to measure ROV implementation status, as in Table 1. Two projects, route views[*5] and RIS[*6], connect to ASes and collect routes to facilitate various measurements. In both cases, the data show that around half to two thirds of ASes have invalid routes when compared with valid routes. However, just because a particular AS does not have invalid routes does not mean that it has implemented ROV. If an upstream AS has implemented ROV, the downstream AS that obtains the routes thus propagated will also no longer have invalid routes. So these results do not indicate that an AS has implemented ROV, but they do demonstrate the effect of ROV in terms of the objective of not propagating invalid routes. These numbers can be expected to change ahead as ROV is increasingly deployed.

2.4 IIJ's Efforts

IIJ is also working on RPKI. FIrstly, at end-2020 IIJ created ROAs for most of the IP addresses it has been allocated by JPNIC (IPv4 82%, IPv6 100%). This allows us to mitigate the risk of IIJ's IP addresses being subject to route hijacking via ASes that have implemented ROV. This effect will increase as more ASes implement ROV. In cases where we have not created ROAs, there are either special circumstances that result in incompatibilities with the JPNIC

*4 RIPE NCC, "Routing Certification Beacons" (https://labs.ripe.net/Members/markd/routing-certification-beacons/).
*5 Routeviews, "University of Oregon Route Views Project" (http://www.routeviews.org/routeviews/).
*6 RIPE NCC, "RRC00 -- RIPE-NCC Multihop, Amsterdam, Netherlands -- Peer List" (http://www.ris.ripe.net/peerlist/all.shtml).

system, which issues the ROAs, or some or all of the addresses are advertised by a customer's AS, so we will need to coordinate with them. We intend to resolve these issues in all cases ahead.

The creation of ROAs for IIJ's allocated IP addresses is going well, but looking at all routes for which AS2497 (IIJ) is the origin AS, only around 30% have ROAs. This is due to customers who use IIJ's services receiving address allocations directly from JPNIC (provider-independent addresses) and using AS2497 as the origin AS. ROAs must be created by the organization that was allocated the addresses, not the AS advertising the BGP routes, so in these cases, the customers should be creating the ROAs themselves. And in these cases, IIJ is indeed encouraging its customers to create ROAs.

IIJ continues to implement ROV on its connections with other ASes and had done this on over 50% of such connections as of end-2020. Connections between ASes can generally be put into three categories: peer connections, upstream (or transit) connections to upstream ISPs, and customer connections whereby the AS provides connectivity to its customers. ROV is implemented on all of IIJ's peer and upstream connections. We have not yet implemented it for customers who purchase connectivity services from IIJ, but we use strict route filtering on points of connection with our customers and thus almost no invalid routes enter the mix. As a result, there are almost no invalid routes within IIJ's network, but even so, implementing ROV for customer connections as well will allow us to more reliably exclude invalid routes, and we thus plan to implement ROV for our customers possibly as early as FY2021.

Customer understanding and cooperation is essential to implementing RPKI for service users as well, but awareness of the importance and need for RPKI remains inadequate. We believe RPKI will be essential to improving not only the stability of customers' data communications but the stability of the Internet of a whole as well, so we are working to raise awareness about RPKI through a range of channels.

2.5 Looking Aheadn

We have discussed origin AS validation using RPKI, but this is not a panacea for all the various sorts of routing failures that occur on the Internet daily. As explained, origin AS validation only involves validating the combination of IP address and origin AS. It cannot detect route hijacking when the origin AS itself is spoofed.

Alongside route hijacking, another problem that frequently occurs is route leaking. This phenomenon, which tends to be due to configuration errors, occurs when routes received from a given AS are propagated by being advertised to other ASes when they shouldn't be. When this happens, traffic passes through ASes that it normally shouldn't, resulting in problems such as substantial traffic delays and packet losses. These incidents actually do occur several times a year on the Internet, affecting prominent online services and ISPs and causing disruptions with a large enough impact to make the mainstream news. Origin AS validation is ineffective against route leaks.

Various technologies and mechanisms for dealing with such events are being studied and discussed, and some are moving toward being standardized and implemented, but they will likely take quite some time to gain full traction given that the idea for RPKI appeared before 2000 and is only now finally beginning to take hold. Even so, now that the Internet has become a key part of our social infrastructure, major failures could have an immeasurable impact. So every AS that makes up part of the Internet should be working hard and consistently to address this, and as a member of the Internet community, IIJ is also doing its utmost in this area.

Takafusa Hori
Manager, Network Technology Section, Network Technology Department, Infrastructure Engineering Division, IIJ
Mr. Hori is engaged in running the IIJ backbone network.