Network Working Group
Internet-Draft
Intended status: Informational
Expires: March 3, 2015

F. Mejia, Ed.
AEPROVI
R. Gagliano
A. Retana
Cisco Systems
C. Martinez
G. Rada
LACNIC

August 30, 2014

Implementing RPKI-based origin validation one country at a time. The Ecuadorian case study. draft-fmejia-opsec-origin-a-country-01

Abstract

One possible deployment strategy for BGP origin validation based on the Resource Public Key Infrastructure (RPKI) is the construction of islands of trust. This document describes the authors' experience deploying and maintaining a BGP origin validation island of trust in Ecuador.

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 http://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 March 3, 2015.

Copyright Notice

Table of Contents

1. Introduction

BGP origin validation based on RPKI [RFC6811] is in early stages of deployment. As with other new technologies, there are impediments to its global adoption as its full value is not yet perceived. Particularly, RPKI based origin validation involves on one side the creation of a large enough set of signed objects and on the other side the application of network policies based on these signed objects by network operators. An operator that does not see a large enough set of signed objects in the RPKI repository system is not encouraged to implement these set of policies. Conversely, IP address space resource holders that are not required by network operators (i.e. transit providers, peers or operators community in general) to create and mantain their RPKI objects have little incentive to do so.

To overcome this bootstrap problem, it is necessary to create a success story that brings enough value to both: network operators and resource holders. Moreover, one possible strategy for the adoption of a security technology is the creation of islands of trust where the technology is fully deployed in a reduced environment. In this direction, some organizations carried forward a full implementation of an island of trust in Ecuador. This was a multi stakeholder project where each party (resource holders, an Internet Exchange Point manager, a Regional Internet Registry and an equipment manufacturer) contributed to its success.

This document describes the experience of implementing RPKI-based origin validation in Ecuador and it is expected to be an useful guide to start other similar projects.

Below, it is described the different roles in the project and the involved parties.

1.1. Policer Network

In this document, the "Policer Network" is the networking infrastructure where the origin validation based on RPKI will be deployed to apply polices on BGP announcements. NAP.EC (www.nap.ec) was selected for this role.

NAP.EC is the Internet Exchange Point (IXP) in Ecuador with two Points of Presence (POPs): Quito (UIO) and Guayaquil (GYE). It has a BGP route-server in each location with a mandatory multilateral routing policy (i.e. all participants have a BGP session to the route-server). Each location uses a different IP address block and Autonomous System Number (ASN). NAP.EC is a meeting point where many organizations (Internet Service Providers, content providers, root servers, etc.) exchange routing information.

The participants connected to NAP.EC announce almost 100% of the total address space used in Ecuador (be believe it is 100% but we cannot be certain though). In some cases they announce their own address space and in some cases they are transit providers for their customers' resources.

AEPROVI (www.aeprovi.org.ec) manages the NAP.EC infrastructure. It is a non-profit organization, based on membership and brings together around 30 Ecuadorian ICT-related companies. AEPROVI also has an excellent reputation as an innovator in the local networking community thanks to projects such as IPv6 adoption and CDN cache servers hosting. These projects have given the local community concrete value and have build the trust on the team that manages the local IXP.

Thanks to this trust, and to the fact that all local BGP announcements are performed through the route servers, NAP.EC is uniquely positioned to become the "policer network" for this project. At the same time, it can be said that implementing origin validation at the NAP.EC route servers is equivalent to implementing it for all inter-domain routing in the country.

1.2. The resource holders

In this document, an organization which operates their own IP prefixes is called resource holder or simply the holder. They may have resources allocated/assigned from a Regional Internet Registry (RIR) and/or legacy resources (if the allocation was done before RIR formation). Resource holders are responsible for creating the RPKI signed objects for this project.

NAP.EC routing tables involve a number of holders, including organizations like Internet Service Providers, content providers, universities, .ec domain administrator and root servers. Most of them are Ecuadorian companies and have received IP resources only from LACNIC, but some have both RIR and legacy resources. Moreover, a few holders are foreign companies and their resources are legacy or from other RIRs (e.g. root servers and content providers).

Not all resource holders are directly connected to the NAP.EC fabric; some have IP address resources but not an ASN and some others are small networks that receive traffic from other bigger networks. In this case their IP address prefixes are announced by their transit providers. One of the main challenges for this project was to identify all the resource holders that needed to be contacted and to encourage network administrators from these organizations to participate.

In addition, some resource holders are part of a larger (and sometimes international) organization, with strong change management processes. This means that any change on their configurations needed to be planned ahead of time and consulted outside of the country.

In NAP.EC - UIO, the routing table includes prefixes used in Ecuador and other countries.

In NAP.EC - GYE, the routing table includes prefixes from companies operating only in Ecuador.

For the project, the target was limited to prefixes used in Ecuador by Ecuadorian holders that had received resources from LACNIC until mid-2013.

1.3. RPKI certificate authorities and repository

The five Regional Internet Registries (RIRs) have a critical role in the RPKI trust model since they manage the trust anchors of the RPKI hierarchic design. Additionally, due to some reasons (e.g. economics, skills) the scenario where the Certification Authority (CA) certificate is hosted by a RIR will be the most popular for a long time, in which case, RIR's online software tools to manage RPKI objects are imperative.

The RIR-hosted RPKI CA model was used for this project. Local RPKI validation servers (validation and cache) were locally deployed. This means that all resource holders had to create and manage their RPKI signed objects using the online tools implemented by LACNIC and that the local validation servers retrieve these objects from the RIR's public global repositories. No local RPKI CA nor repository were configured.

LACNIC also runs a RPKI testbed (test CA with correspondent GUI and Trust Anchor material). This infrastructure was used during the training activity.

1.4. The technical support

RPKI and origin validation are in the early stage of deployment. Few people have full knowledge about its RFCs, the implementation support in different routers and the maintenance of RPKI signed objects. To involve trained people and train new ones is very important.

People from an equipment manufacturer (Cisco) contributed with support in the startup stage and to train the holders' staff. LACNIC's staff contributed developing new online RPKI tools and training about how to use them.

2. Objective

Considering all the definitions given during the introduction and after several discussions through face and online meetings among the involved parties, the following objective was agreed on:

"Deploy RPKI-based BGP origin validation in NAP.EC's route servers. For the success of the project, 80% of the Ecuadorian prefixes (both IPv4 and IPv6) received by those routers should have a valid origin."

In order to monitor the progress, NAP.EC - GYE was taken as reference because NAP.EC - UIO had non-Ecuadorian prefixes announced.

3. Planning

The project started with an initial idea from a very reduced number of enthusiasts that identified a suitable network (the island of trust), involved the appropriate organizations and set milestones in order to carry forward a full implementation of the technology. Into the process, all parties identified the gaps and proposed solutions to overcome them.

One point that it was wanted to guarantee is that we would be able to create the appropriate "buzz" around the project. So, a communication strategy should not be overlooked. In this case, LACNIC and AEPROVI signed a MoU in April 2013 and all parties (LACNIC, AEPROVI and Cisco) announced the project and issued a press release at the LACNIC event in May 2013.

Some points that required specific discussion by the core team included:

1. RPKI-based origin validation support in the route-servers equipments

2. How to deploy a RPKI cache into the Network

3. How to populate the RPKI database with the correct and necessary information

4. Action to take with NotFound and Invalid prefixes

3.1. RPKI-based origin validation support

NAP.EC uses Cisco equipment. The project started with the initial idea of to implement origin validation into existing routers used as route servers, simply after a software update or upgrade. However, the vendor had no plans to support it in the existing platform. AEPROVI had future plans to carry forward a routers renewal, then this issue was overcame but it stopped the project for some time. Describing the equipment renewal process is beyond the scope of this document.

For Cisco equipment, the vendor has made available some online software tools to check the support. About origin validation, the routers must support: RTR protocol [RFC6810] and RPKI-based origin validation [RFC6811]. Moreover, among other things, four octects ASN support [RFC6793] and IPv6 routing support ([RFC2545] and some others) are mandatory in NAP.EC.

The selected routers were two Cisco ASR-1000 series routers (one for Quito and other for Guayaquil).

3.2. Deploying a RPKI cache into the network

Based on available resources and existing skills, it was decided to use Virtual Machines (VM) as RPKI caches, which would run GNU Linux.

The validating software is in the early stages of its development and there could be bugs or reliability problems, so it was decided using two different packages (processes) in each VM.

To ensure high availability, it was decided to deploy two VMs, each one in different host server.

There are no servers in Guayaquil, therefore both VMs would be in Quito within the NAP.EC management network and connect via the RTR protocol to the route-servers located in Quito and Guayaquil.

Additionally, the firewall rules allow RTR connections from the NAP.EC LANs to the RPKI validator servers in order to facilitate participants to perform origin validation in their edge equipments (if they wish to in the future).

3.3. Populating the RPKI database

The IP resource holders must create all needed RPKI data for the project, at least certificates and Route Origin Authorizations (ROAs). Moreover, the technical staff needed training about RPKI and origin validation because it is a new technology. Accordingly, a reasonable method to achieve it should be contrived.

It was decided to organize an event with two objectives: training and RPKI object signing. One key planning activity was to create the list of participants and to make sure that at least one participant per network had the authentication credentials to the LACNIC system to create its RPKI signed objects.

The target community was limited to Ecuadorian organizations that had received IP resources from LACNIC until mid-2013. That meant around fifty (50) organizations including Internet Services Providers, universities, banks, etc., or expressing it in prefixes: around 8600 IPv4 and 60 IPv6 blocks.

Some weeks before the deployment, there was informal dissemination meetings between the NAP.EC administrator and the participants. The project milestones were reported and the attendees received information about RPKI and origin validation for the first time. A

complete training was offered as a project milestone in the next few weeks.

All organizations were contacted and received an invitation to the event. More information about this can be found in Section 5.

3.4. Action to take with NotFound and Invalid prefixes

Despite the efforts, the RPKI database information may be incomplete, therefore the routing tables often will have NotFound prefixes. Moreover, it is needed some time after the first contact with RPKI-based origin validation technology to fix possible errors (e.g. invalid prefixes) and to assess the impact. A strict policy of dropping prefixes did not seem convenient as a starting point for the project.

It was decided that NAP.EC proceeds as follows:

- At the beginning, the NAP.EC's routers only would monitor the RPKI origin state of prefixes without action.

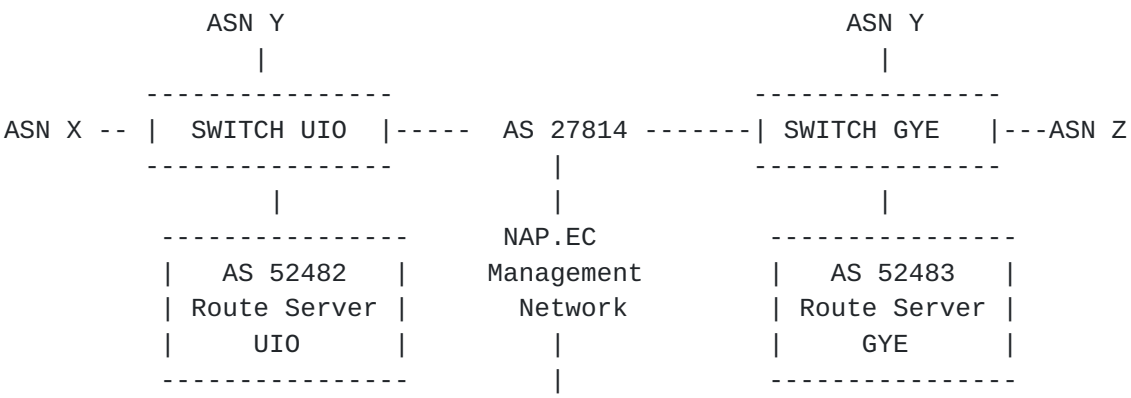
- In the near future, NAP.EC administration might change the action based on results of the signing-party event and community consensus.

As part of the second stage, some days after the signing event, each prefix is being marked with a BGP community to identify its RPKI origin state, sending that information to the participants.

Finally, some months later, a date was set to begin applying a strict policy. The policy was defined as follows: dropping Invalid prefixes and setting a lower local preference for NotFound prefixes.

4. Deployment

Following is the NAP.EC topology during the deployment:



RPKI Validation servers
(2 virtual server at UIO)

4.1. RPKI Validation servers

A virtual machine (named VM1) on VMware ESXi was deployed, running GNU Linux, Centos distribution. The other one (named VM2) was cloned from this one.

Each virtual machine has access to the Internet through 1 (one) ethernet interface with a public IPv4 address within the same subnet like this: 192.0.2.2/27 for VM1, 192.0.2.3/27 for VM2 and 192.0.2.1/27 for network gateway.

The 192.0.2.0/27 network is within AS 27814. AS 27814 contains NAP.EC monitoring equipment (among other things) and it is connected to NAP.EC - Quito and NAP.EC - Guayaquil.

Each VM has the following packages (installation and configuration guides of these packages are beyond the scope of this document):

- ntpd (as NTP client)
- iptables (as firewall)
- apache (as web server)
- validating software from RIPE (http://www.ripe.net/lir-services/resource-management/certification/tools-and-resources)
- validating software from the rpki.net project (http://rpki.net/wiki/doc/RPKI/Installation)

Each validating software was setup on a different port:

- validating software from RIPE, port 65001,

- validating software from the rpki.net project, port 65002.

Each validating software has a monitoring web page, each one was configured on different port:

- RIPE software, port 65081,

- rpki.net software, port 65082.

4.2. Origin validation setting

Since the routers renewal, NAP.EC - Quito and NAP.EC - Guayaquil have each one a Cisco ASR-1000 series router as route server. These routers have IOS-XE version 3 which runs a process with IOS version 15.

First, it is required to configure communication via RTR protocol between the routers and the RPKI validation servers (caches). In this step, the necessary data are: IP addresses and service ports of the caches and the time which the router will re-query the cache (refresh time). Currently, RPKI information does not change quickly, therefore 600 seconds (10 minutes) may be considered enough for refresh time.

Cisco IOS 15 drops Invalid prefixes by default, but there is a command to avoid this behavior (ibgp bestpath prefix-validate allow-invalid). This must be applied while the policy is no action.

Later, a route-map is required to configure any action as applying different BGP local preference or marking each prefix with a BGP community based on its RPKI origin state (also send-community option must be enabled within the BGP session configuration):

The following community assignment policy was applied:

- <IXP-ASN>:21 --> Valid origin

- <IXP-ASN>:22 --> NotFound origin

- <IXP-ASN>:23 --> Invalid origin

Where <IXP-ASN> equals:

- 52482 for NAP.EC - Quito, or

- 52483 for NAP.EC - Guayaquil.

The template used in NAP.EC is in Appendix A.

5. Training and RPKI signing event

The event was called "Seminario sobre seguridad en el encaminamiento de Internet: BGP RPKI - Validacion de origen" and was scheduled for September 4-5, 2013. The agenda included theoretical and practical training, plus two time slots to sign RPKI objects: one at the end of the first day and other one during the second day.

Lack of training materials was a issue to overcome during preparatory work of the event. Some necessary activities were:

- The instructors (four people) prepared materials to cover topics such as BGP, RPKI, origin validation and the new NAP.EC platform.

- LACNIC's staff developed two new on-line tools: RPKI ROA wizard (http://tools.labs.lacnic.net/roa-wizard/) and RPKI announcement (http://tools.labs.lacnic.net/announcement/), further improved the demo environment of the RPKI system (http://rpkidemo.labs.lacnic.net/).

- Cisco's staff implemented a temporary virtualized network with many routers supporting RPKI and origin validation.

The event took place in a hotel and had Internet to access the training tools and the real LACNIC's hosted RPKI system.

Not all organizations sent a representative. The attendance represented around 80% of the target prefixes.

6. Outcome and post-event activities

Before the event, less 1% of the Ecuadorian prefixes were signed. At the start of the second day, less than 20% of the Ecuadorian prefixes were covered by a ROA. At the end of the event, around 80% of the Ecuadorian prefixes had a RPKI origin state as Valid.

MRTG graphs were implemented to monitor the amount of Valid, NotFound and Invalid prefixes after the event.

Feedback was received from attendees before closing the event. Some people recommended applying an acceptable policy in order do not waste the successful effort.

A few days after the event, some non-attending organizations were contacted by the NAP.EC administrator and meetings were coordinated for ROA creation. After these activities, almost 100% of Ecuadorian prefixes are covered for a ROA.

Communication activities performed after the event included:

- This document and presentation at relevant IETF Working Groups.
- Presentation at IEPG, LACNIC and other NOG events.
- Publication at tech sites.
- Communication to local telecommunications regulator.
- Document and presentation at CITEL (Organization of American States).
- Blogging and social media in relevant platforms.

As subsequent operational tasks, the following are necessary:

- Updating of validating software.
- Permanent monitoring the origin state of prefixes.
- Alerting about Invalid prefixes.

7. Lessons learned and best practices

- Implementation support needs to be verified in all target platforms.
- The IP resource holders community need RPKI-based origin validation training.
- Initial work to have the "right people" in the room is a key to success for the RPKI signing party. Particularly, operators need to have access to their RIR account.
- One day for a RPKI signing party is insufficient, two days is a better practice. The participants may not be confident about their skills or may need further authorization (people need to sleep over what they learned the first day).
- The event was a great opportunity to assemble the local community, particularly resource holders that had no previous participation at the local IXP.
- Post event communication needs to be discussed ahead of time.

- Operators are less conservative than original though by organizers and once RPKI local space was full, support for removing invalid prefixes was unanimous.

- From now on, when a new ISP wants to join NAP.EC, it receives information about RPKI-based origin validation and it is invited to create its ROAs.

8. IANA Considerations

No IANA requirements

9. Security Considerations

This document describes the experience of implementing a BGP origin validation island of trust in Ecuador. The actions taken are explicitly to be able to validate the origin in a BGP advertisement. There were no security-related issues identified during the deployment.

10. Acknowledgements

The authors wish to thank:

- all attendees at the training and RPKI signing event, without them this would not have happened.

- AEPROVI, LACNIC and Cisco for supporting the project.

- Arturo Servin for supporting the project from the start.

- Francisco Balarezo, Andres Piazza, Nicolas Fiumarelli and Chip Sharp as well as ISOC and Andean-Trade.

11. Informative References

[RFC2545] Marques, P. and F. Dupont, "Use of BGP-4 Multiprotocol Extensions for IPv6 Inter-Domain Routing", RFC 2545, March 1999.

[RFC6793] Vohra, Q. and E. Chen, "BGP Support for Four-Octet Autonomous System (AS) Number Space", RFC 6793, December 2012.

[RFC6810] Bush, R. and R. Austein, "The Resource Public Key Infrastructure (RPKI) to Router Protocol", RFC 6810, January 2013.

[RFC6811] Mohapatra, P., Scudder, J., Ward, D., Bush, R., and R. Austein, "BGP Prefix Origin Validation", RFC 6811, January 2013.

Appendix A. Router configuration templates for Cisco IOS

TEMPLATE 1

Policy: Only marking prefixes based on RPKI origin state.

```
router bgp <IXP-ASN>

   bgp rpki server tcp 192.0.2.2 port 65001 refresh 600

   bgp rpki server tcp 192.0.2.2 port 65002 refresh 600

   bgp rpki server tcp 192.0.2.3 port 65001 refresh 600

   bgp rpki server tcp 192.0.2.3 port 65002 refresh 600

   !

   neighbor <neighbor-IPv4> remote-as <neighbor-IPv4-ASN>

   neighbor <neighbor-IPv4> version 4

   !

   neighbor <neighbor-IPv6> remote-as <neighbor-IPv6-ASN>

   neighbor <neighbor-IPv6> version 4

   !

   address-family ipv4

      bgp bestpath prefix-validate allow-invalid

      neighbor <neighbor-IPv4> send-community

      neighbor <neighbor-IPv4> route-map <route-map-name> in

   exit-address-family

   !

   address-family ipv6
```

```
            bgp bestpath prefix-validate allow-invalid
            neighbor <neighbor-IPv6> send-community
            neighbor <neighbor-IPv6> route-map <route-map-name> in
         exit-address-family
      !
      !
      ip bgp-community new-format
      !
      !
      route-map <route-map-name> permit 10
         match rpki valid
         set community <IXP-ASN>:21
      !
      route-map <route-map-name> permit 20
         match rpki not-found
         set community <IXP-ASN>:22
      !
      route-map <route-map-name> permit 30
         match rpki invalid
         set community <IXP-ASN>:23
      !
```

TEMPLATE 2

Policy: Dropping Invalid prefixes and setting lower local preference for NotFound prefixes.

```
   router bgp <IXP-ASN>
```

```
      bgp rpki server tcp 192.0.2.2 port 65001 refresh 600
      bgp rpki server tcp 192.0.2.2 port 65002 refresh 600
      bgp rpki server tcp 192.0.2.3 port 65001 refresh 600
      bgp rpki server tcp 192.0.2.3 port 65002 refresh 600
      !
      neighbor <neighbor-IPv4> remote-as <neighbor-IPv4-ASN>
      neighbor <neighbor-IPv4> version 4
      !
      neighbor <neighbor-IPv6> remote-as <neighbor-IPv6-ASN>
      neighbor <neighbor-IPv6> version 4
      !
      address-family ipv4
         neighbor <neighbor-IPv4> send-community
         neighbor <neighbor-IPv4> route-map <route-map-name> in
      exit-address-family
      !
      address-family ipv6
         neighbor <neighbor-IPv6> send-community
         neighbor <neighbor-IPv6> route-map <route-map-name> in
      exit-address-family
   !
   !
   ip bgp-community new-format
   !
```

```
      !
      route-map <route-map-name> permit 10
         match rpki valid
         set community <IXP-ASN>:21
      !
      route-map <route-map-name> permit 20
         match rpki not-found
         set local-preference 50
         set community <IXP-ASN>:22
      !
      !
```

Authors' Addresses

Fabian Mejia (editor)
AEPROVI
Av. Republica de El Salvador N34-211
Quito
EC

Email: fabian@aeprovi.org.ec

Roque Gagliano
Cisco Systems
Avenue des Uttins 5
Rolle 1180
Switzerland

Email: rogaglia@cisco.com

Alvaro Retana
Cisco Systems
7025 Kit Creek Rd.
Research Triangle Park, NC 27617
US

Email: aretana@cisco.com

Carlos Martinez
LACNIC

Email: carlos@lacnic.net

Gerardo Rada
LACNIC