Secure Interdomain Traffic Exchange

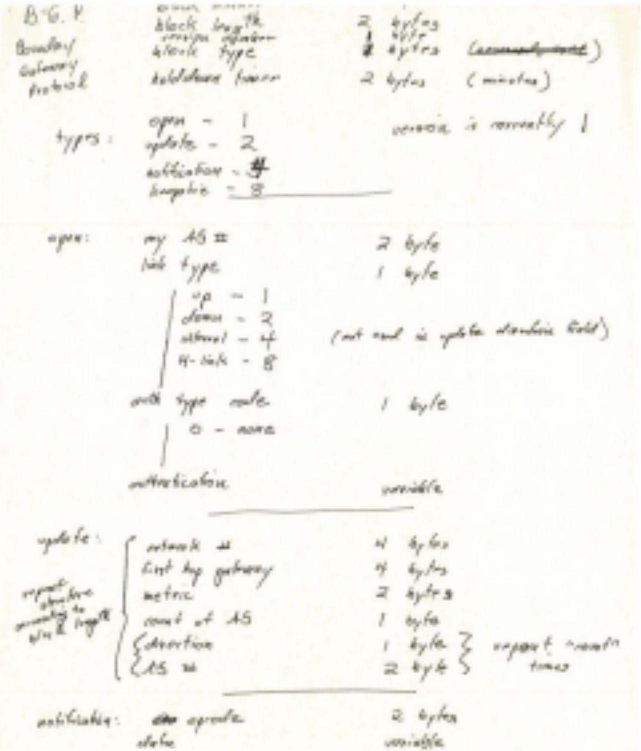
iNOG::13

June, 6 2019

Sergey Kolobov

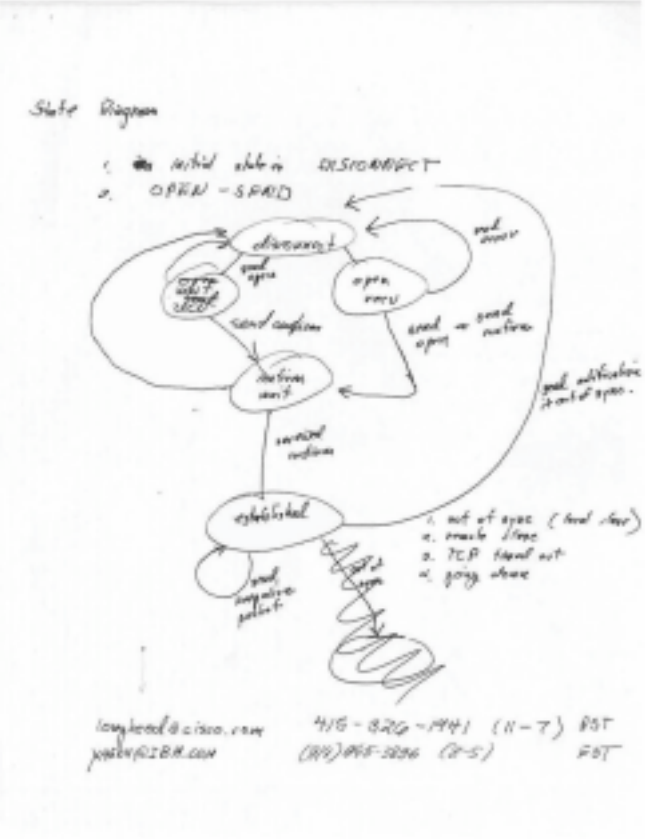
How it all started...

It was 1989...



B.G.P.
Boundary Gateway Protocol

	block length	2 bytes	
	version number	1 byte	
	block type	1 bytes	~~(command/event)~~
	holddown timer	2 bytes	(minutes)
types:	open - 1 update - 2 notification - ~~3~~ 4 keepalive - 8		version is currently 1
open:	my AS #	2 byte	
	link type	1 byte	
	up - 1 down - 2 internal - 4 H-link - 8		(not used in update direction field)
	auth type code	1 byte	
	0 - none		
	authentication	variable	
update:	network #	4 bytes	
repeat structure according to block length	first hop gateway	4 bytes	
	metric	2 bytes	
	count of AS	1 byte	
	direction	1 byte	repeat count times
	AS #	2 byte	
notification:	~~the~~ opcode	2 bytes	
	data	variable	

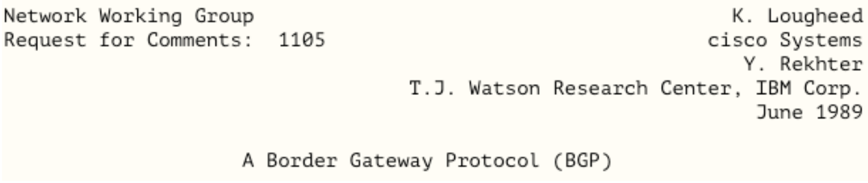


State Diagram

1. ~~the~~ initial state is DISCONNECT
2. OPEN - SEND

1. out of sync (total clear)
2. create close
3. TCP timed out
4. going down

lougheed@cisco.com 415-326-1941 (11-7) PST
yakov@IBM.com (914)945-3896 (8-5) EST



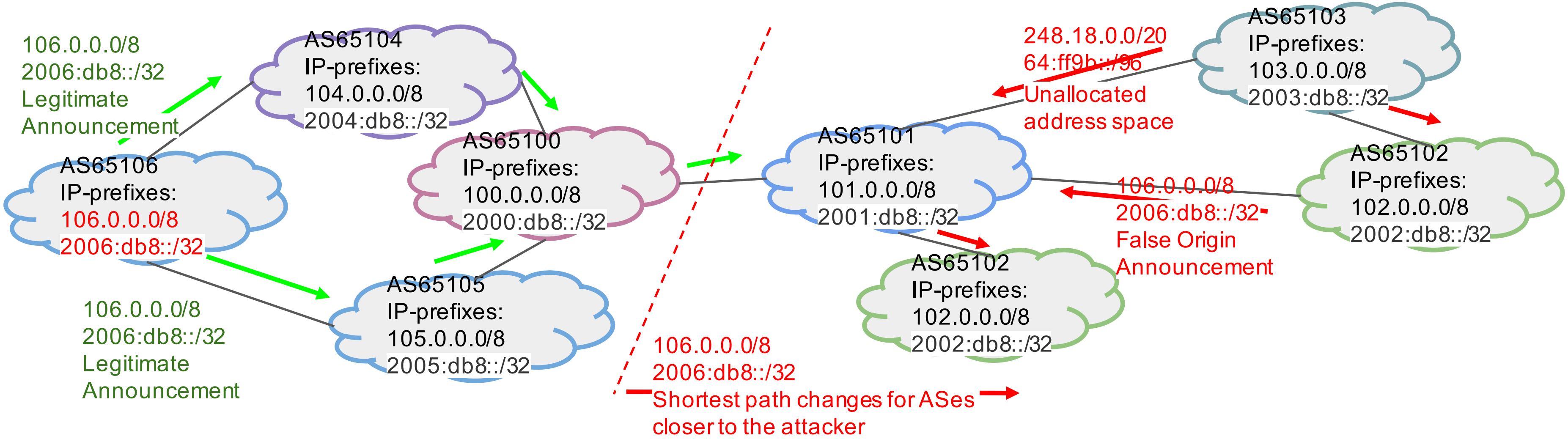
```
Network Working Group                                         K. Lougheed
Request for Comments:  1105                                 cisco Systems
                                                               Y. Rekhter
                                   T.J. Watson Research Center, IBM Corp.
                                                                June 1989

                   A Border Gateway Protocol (BGP)
```

BGP incidents

- Nov’18 - Google prefixes leaked, traffic redirected to China Telecom for 74 minutes
  - AS37282 MainOne, Nigerian ISP
- Apr’18 - myetherwallet.com
  - AS10297 eNET
- Jul’18 - Instagram rerouted to Iran
  - AS 58224 Iran Telecom PJS
- Dec’17 - Google, Apple, NTT, Facebook, Riot Games…
  - AS 39523 DV-LINK-AS
- Apr’17 - VISA, MasterCard, Symantec…
  - AS 12389 Rostelecom
- Jul’17 - Savvis, Century Link, Mercury Payment Systems
  - AS 38146 Digital Wireless Indonesia, AS 38182 Extreme Broadband
- Apr’17 - AWS Route53, MyEtherwallet.com
  - AS 10297 eNet, OH, USA
- 2014 - Canadian Bitcoin Exchange Hijack
- 2008 - YouTube Hijack
  - AS17557 Pakistan Telecom

BGP Control Plane Attacks

- Prefix Hijacking - AS origin prefix that is not authorized by prefix owner
- Sub-Prefix Hijacking - AS announces a more specific (longer) prefix that the owner
- Prefix Squatting - AS origin allocated but unused address space
- AS Path modification - AS removes some of the preceding ASes in AS_PATH to make it look shorter and Tx it
- Kapela-Pilosov Attack - AS replaces a prefix in a Rx update by a more specific prefix and Tx it
- Route Leaking - Dual-home Stub/Customer AS leaks route to upstream ISP1 about routes in upstream ISP2



Best practices

- MANRS - Mutually Agreed Norms for Routing Security
  - is a global initiative, supported by the Internet Society, that provides crucial fixes to reduce the most common routing threats.
  - Network Operators, IXP, Enterprise, Service Providers
  - Filtering
  - Anti-Spoofing
  - Coordination
  - Global Validation
  - Prevent Propagation
  - Protect Peering Platform
  - Facilitate ISP Communication
  - Provide Monitoring Tools

- NIST special publication on Secure Interdomain traffic exchange
  - National Institute of Standards and Technology, U.S. Dept. of Commerce
  - Technologies recommended in this document for securing the interdomain routing control traffic

NEW

Route Leak Detection and Filtering using Roles in Update and Open messages

- Avoiding Route Leaks Optional non-transit attribute
  – Internal Only To Customer Attribute (iOTC):
    - Flag is not set – announce in all directions
    - Flag is set – announce only to internal and customer links

- Detecting Route Leaks Optional transitive attribute
  – External Only To Customer Attribute (eOTC):
    - Attribute is not set – no info
    - Attribute is set and equals to neighbor AS – ok
    - Otherwise – route leak

  - Can’t filter based on this as MitM can change attribute and affect reachability to the victim

Optional attributes and Communities aren’t solving the problem, need something else, obligatory:

- BGP Role is new configuration option that SHOULD be configured on each BGP session based on BGP capability in UPDATE and OPEN message

https://tools.ietf.org/html/draft-ietf-idr-bgp-open-policy-05

https://tools.ietf.org/html/draft-ymbk-idr-bgp-eotr-policy-02

soBGP

- Validate an AS is authorized to originate a prefix.
- Verify a peer which is advertising a prefix has at least one valid path to the destination.

- ISP X publishes information about its connections;
- ISP Y publishes information about its connections;

If there are both pairs (X,Y) && (Y,X) – the pair becomes trustable!
If there is only one pair (X,Y) || (Y, X) the pair becomes… less trustable!

- Problems with IXes
- The two side adjacencies don't provide a way to automate anomaly detection without high adoption rate - an attacker can easily up a one-way adjacency.

NEW

ASPA - Autonomous System Provider Authorization

This procedure uses a shared signed database of customer-to-provider relationships that is built using a new RPKI object - Autonomous System Provider Authorization (ASPA).

ASPAs are digitally signed objects that attest that a Customer AS holder (CAS) has authorized a particular Provider AS (PAS) to propagate the Customer's IPv4 or IPv6 BGP route announcements onwards—to the Provider's upstreams or peers.

If valid route is received from customer or peer it MUST have only customer-to-provider pairs in its AS_PATH.
Then if we have a validated database of customer-to-provider pairs we will be able to verify routes received from customers and providers!

```
ASPA :=     {
      customer_asn – signer
      provider_asn – authorized to send routes to upper providers or peers

      AFI – IPv4 or IPv6
            }
```

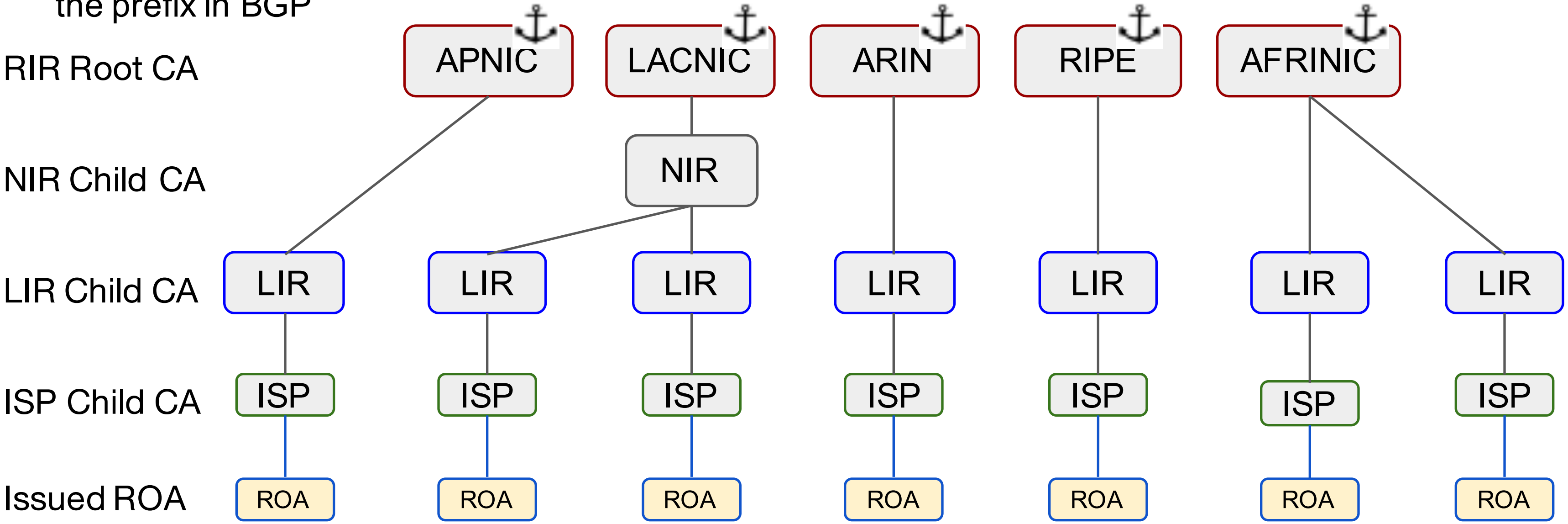
- Do not support announces from provider to client, so hijacks are still possible

https://datatracker.ietf.org/doc/draft-azimov-sidrops-aspa-verification/

RPKI: The chain of trust

Solving the problem of IRR DBs via authoritative, cryptographically verifiable statements by any legitimate IP resource holders:

- X.509 (RFC5280) w/ extension for IP-address and AS ID (RFC3779)
- Trust anchors: RIRs - Route Origin Validation a stepping stone to Path Validation
- Mimic the purpose of Route Objects in IRR
- Route Origin Authorisation (ROA) - signed statement about which AS is authorised to originate the prefix in BGP



RIR Root CA: APNIC, LACNIC, ARIN, RIPE, AFRINIC

NIR Child CA: NIR

LIR Child CA: LIR, LIR, LIR, LIR, LIR, LIR, LIR

ISP Child CA: ISP, ISP, ISP, ISP, ISP, ISP, ISP

Issued ROA: ROA, ROA, ROA, ROA, ROA, ROA, ROA

RPKI: Challenges

ROA validation can’t be used to:

- filter route leaks - if used alone
- filter malicious hijacks - prepend “valid” source AS to forged route announcement

Hierarchical dependency

- if upstream authority didn’t obtain certificates for allocated address space

Roughly a third of records in RPKI are erroneous - principle “do not harm” is under threat

ROA Propagation Time - up to 8 hours to receive new ROAs

Chicken and egg:

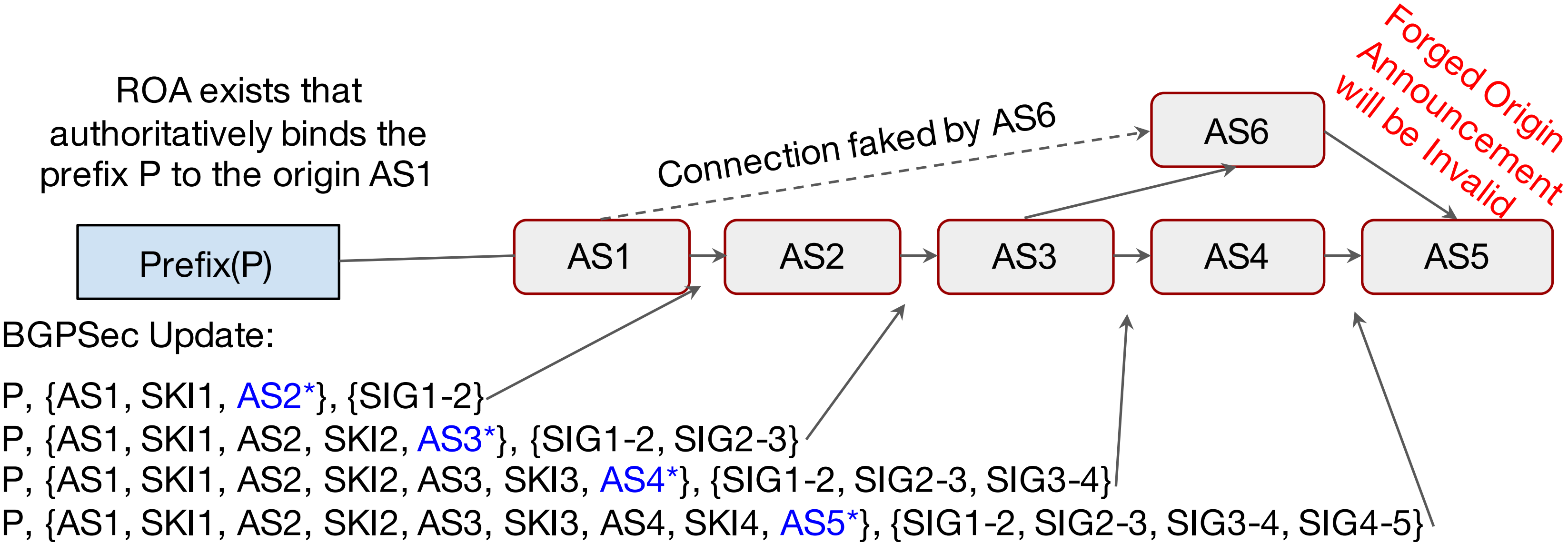
- Certification is effective only when ROA Validation is deployed
- ROA Validation is effective only WRT certified IP-address blocks

Multiple cache servers stale state of a ROA - RFC8481

- validating every prefix regardless of where it comes from, even if local originated and just sets a flag that you can match on in a route map or where ever
- Don’t take any other default actions

RPKI: Challenges -> BGPSec a.k.a. SIDR

RPKI provides only origin validation, AS path validation is specified in **BGPSec** [RFC8205]



ROA exists that authoritatively binds the prefix P to the origin AS1

BGPSec Update:

P, {AS1, SKI1, AS2*}, {SIG1-2}
P, {AS1, SKI1, AS2, SKI2, AS3*}, {SIG1-2, SIG2-3}
P, {AS1, SKI1, AS2, SKI2, AS3, SKI3, AS4*}, {SIG1-2, SIG2-3, SIG3-4}
P, {AS1, SKI1, AS2, SKI2, AS3, SKI3, AS4, SKI4, AS5*}, {SIG1-2, SIG2-3, SIG3-4, SIG4-5}

*Next Hop AS is signed over but not included in the fwd BGPSec Update

Authenticates the entire AS path, from the origin to the traffic source.

NEW

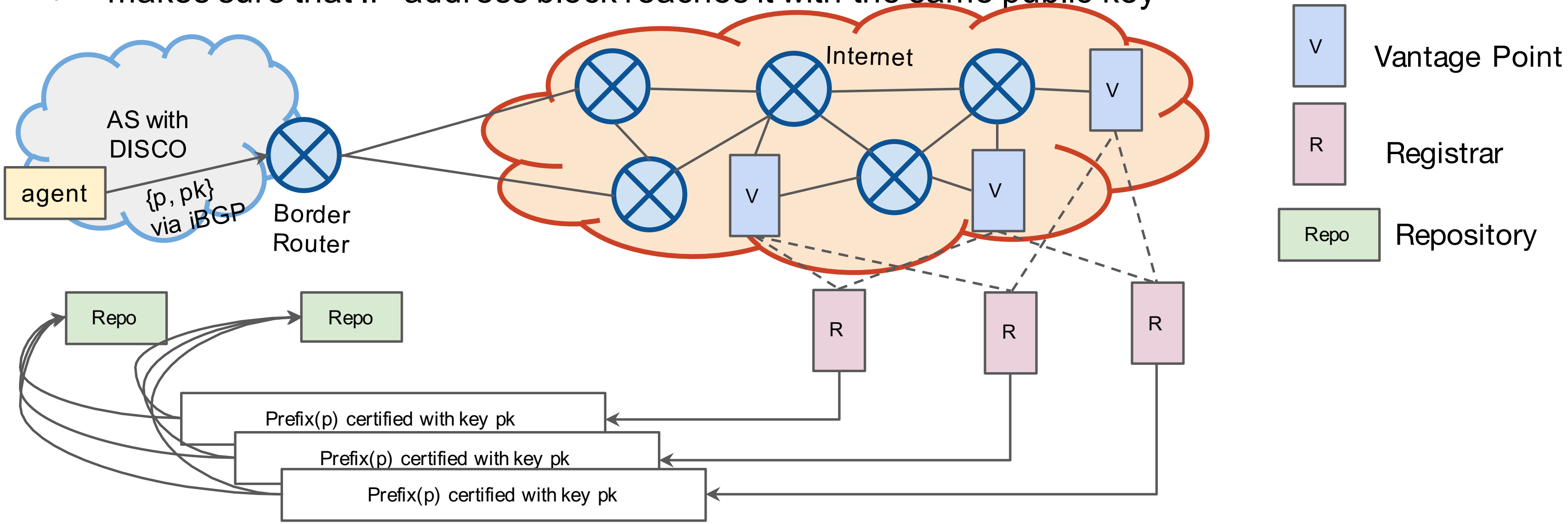
DISCO: High Level Overview

Agent - ensures owner’s AS public key is attached to BGP prefix announcements

- optional transitive attribute
- 256-bit public key

Registrar - monitors BGP advertisements from vantage points,

- approves the ownership of an IP during the “certification time interval” = 1 week
- makes sure that IP-address block reaches it with the same public key



http://www.cs.huji.ac.il/~schapiram/DISCO_HotNets.pdf

Conclusion

Main Challenges

- DISCO - the main challenge is the ability of an attacker to sabotage a prefix certification
- RPKI - Resource DB consistency and data discrepancies
- Legal challenges will apply to both DISCO and RPKI
- BGPSec - the cost of computation

Steps ahead

- Every AS has to utilise up-to-date filters
- Operators must keep IRRdb with up-to-date information
- Push Tier 1 ISPs, Internet Giants and Major IXPs to deploy RPKI
- 80:20 Rule - at the end of the day not everyone is expected to participate
- Seriously consider our options on AS path validation - updates include the gateway, gw can be probed. Found that is unreachable, so it won't install.

Deploy RPKI - it’s an imperative!

Conclusion

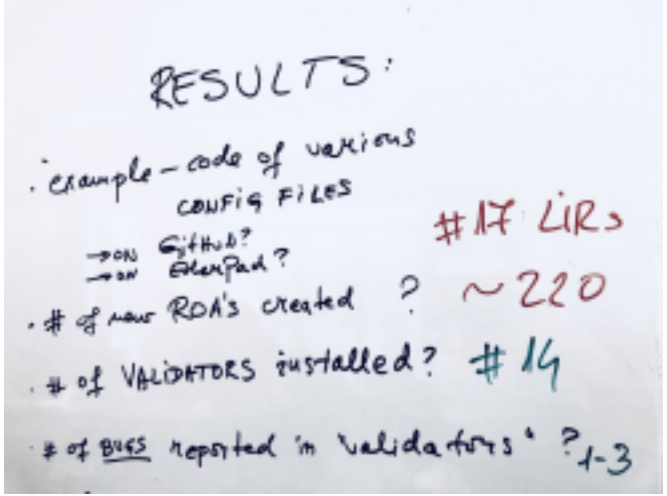
“The AT&T/as7018 network is now dropping all RPKI-invalid route announcements that we receive from our peers.”

“Liquid Telecom, SEACOM and Workonline represent more than two-thirds of Africa's Internet, we have all agreed to turn on RPKI Origin Validation on our networks on 1st April, 2019.”

The Calgary Internet Exchange (YYCIX), MSK-IX, DE-IX, AMS-IX and others using RPKI

RIPE NCC Deployathon, Amsterdam, 7-8 March 2019

RIPE NCC IRR Database Non-Authoritative Route Object Clean-up



RESULTS:

- example-code of various CONFIG FILES
  - → ON GitHub?
  - → ON EtherPad?
- # of new ROA's created ? ~220
- # of VALIDATORS installed? #14
- # of BUGS reported in validators' ? 1-3

#17 LIRs

What you can do...see if your prefixes are signed

```
$ whois -h  whois.bgpmon.net 200.7.86.0
Prefix:              200.7.86.0/24
Origin AS:           28001
Origin AS Name:      LACNIC - Latin American and Caribbean IP address, UY
RPKI status:         ROA validation successful
```

```
$ whois -h  whois.bgpmon.net " --roa 28001 200.7.86.0/24"
0 - Valid
Origin ASN:       AS28001
Not valid  Before: 2017-04-28 04:00:00
Not valid  After:  2023-04-28 04:00:00  Expires in 3y328d18h50m37.7999999821186s
Trust Anchor:     repository.lacnic.net
Prefixes:         200.3.12.0/22 (max length /24)
              200.10.60.0/23 (max length /24)
              200.7.86.0/24 (max length /24)
              2001:13c7:7012::/47 (max length /47)
              2001:13c7:7010::/46 (max length /47)
              2001:13c7:7002::/48 (max length /48)
```

GUI:

http://localcert.ripe.net:8088/roas or https://bgpview.io/

If not - request your ISP/LIR to do so

Something to read:

- RPKI Documentation - https://rpki.readthedocs.io/en/latest/index.html
- NIST Special Publication - https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-189-draft.pdf
- DISCO - https://www.researchgate.net/publication/328896238_Perfect_is_the_Enemy_of_Good_Setting_Realistic_Goals_for_BGP_Security
- An Infrastructure to Support Secure Internet Routing - https://tools.ietf.org/html/rfc6480
- NANOG74 Security Track - https://pc.nanog.org/static/published/meetings/NANOG74/1760/20181003_Tzvetanov_Security_Track_Bgp_v1.pdf
- BGP Prefix Origin Validation - https://tools.ietf.org/html/rfc6811
- The Resource Public Key Infrastructure (RPKI) to Router Protocol - https://tools.ietf.org/html/rfc6810
- Signaling Prefix Origin Validation Results from a Route Server to Peers - https://tools.ietf.org/html/draft-ietf-sidrops-route-server-rpki-light-02
- BGP Prefix Origin Validation State Extended Community - https://tools.ietf.org/html/rfc8097
- RPKI - The required cryptographic upgrade to BGP routing - https://blog.cloudflare.com/rpki/
- Bamboozling Certificate Authorities with BGP - https://www.usenix.org/conference/usenixsecurity18/presentation/birge-lee
- Mutually Agreed Norms for Routing Security (MANRS) Implementation Guide - https://www.manrs.org/wp-content/uploads/2018/03/MANRS-BCOP-20170125.pdf
- Will the SIDR model succeed where the IRR model failed? (Part I) - https://blog.apnic.net/2015/06/01/will-the-sidr-model-succeed-where-the-irr-model-failed-part-i/
- BGPsec and Reality - https://rule11.tech/bgpsec-and-reality/
- How Secure are Secure Interdomain Routing Protocols? - http://www.cs.yale.edu/homes/schapira/BGPAttack.pdf
- Verification of AS_PATH Using the Resource Certificate Public Key Infrastructure and Autonomous System Provider Authorization - https://datatracker.ietf.org/doc/draft-azimov-sidrops-aspa-verification/

Thank you!
Questions?