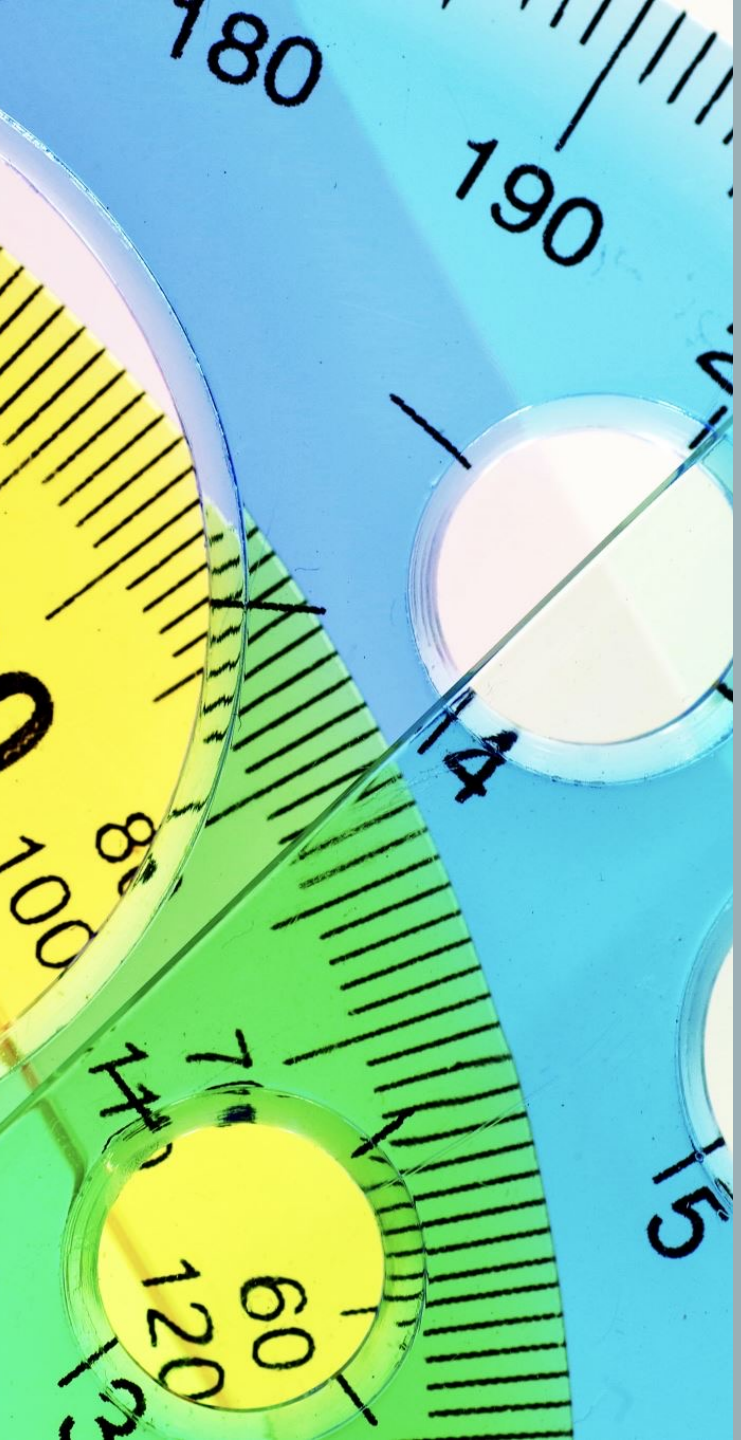


RPKI TIME-OF-FLIGHT

ZANOG-23

2023.03.22

Romain Fontugne, Amreesh Phokeer, Cristel Pelsser, Kevin Vermeulen, & Randy Bush

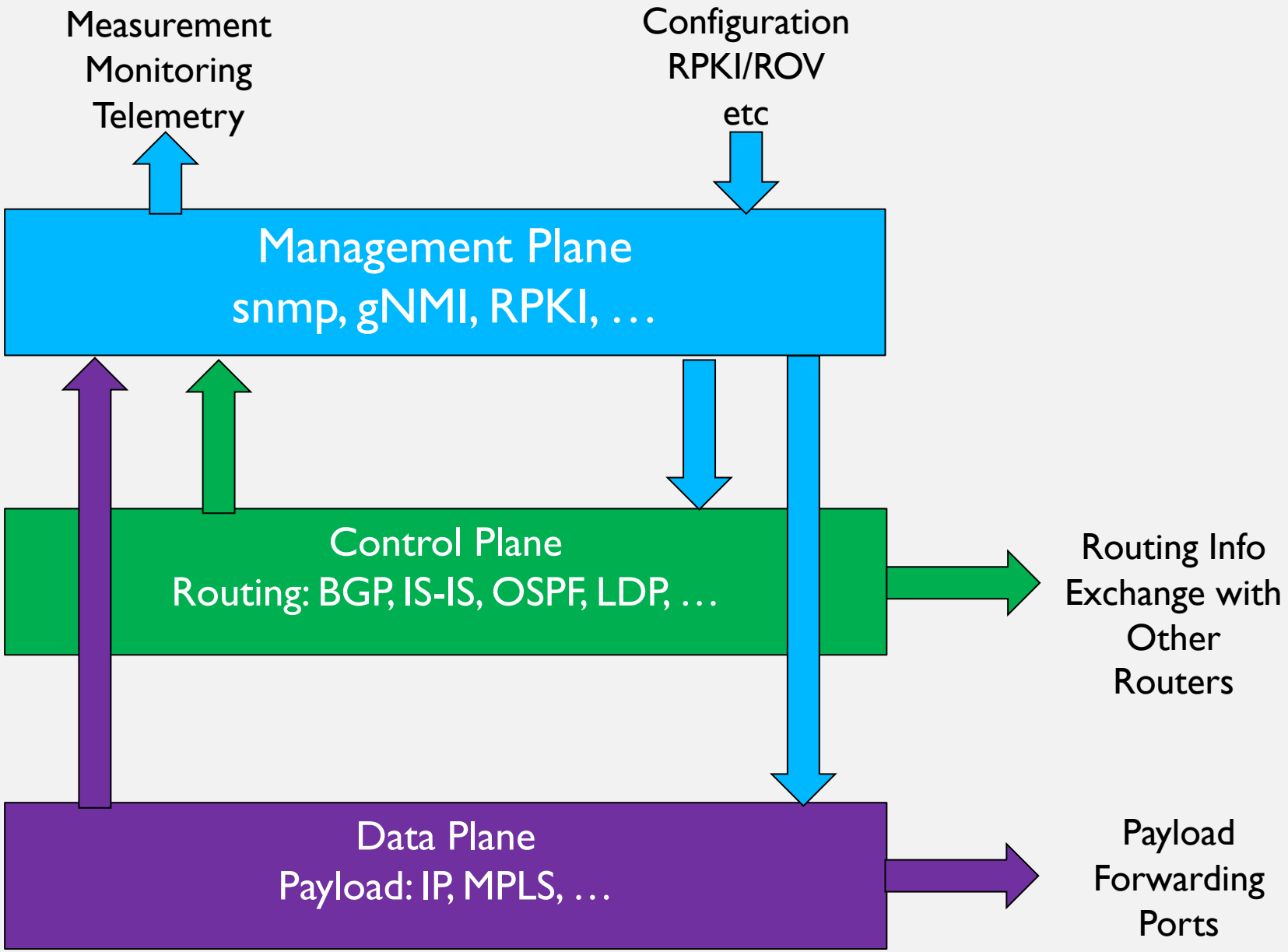
BGP: THE HORROR SYSTEM

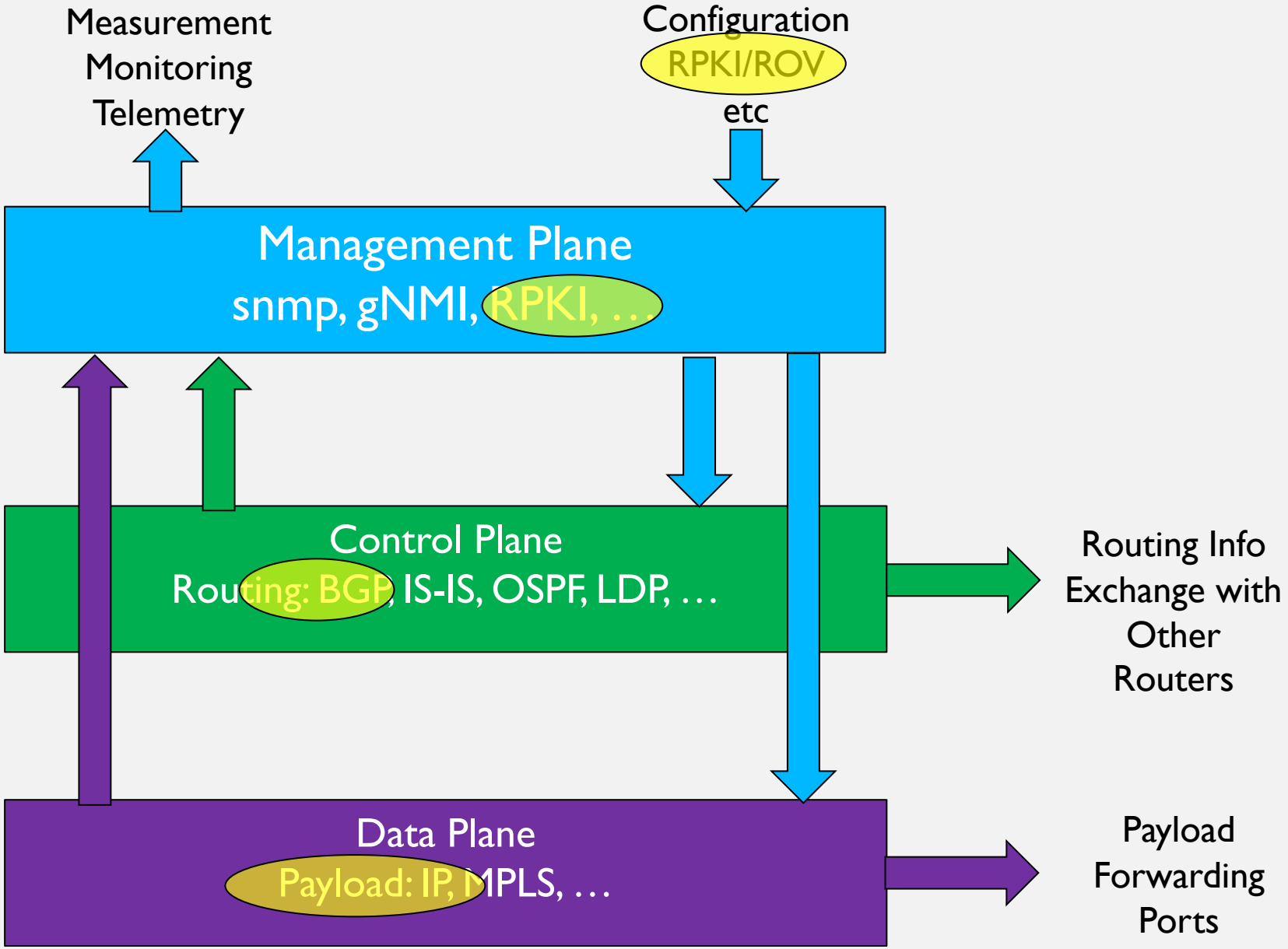
Border Gateway Protocol (BGP) is based entirely on trust between networks

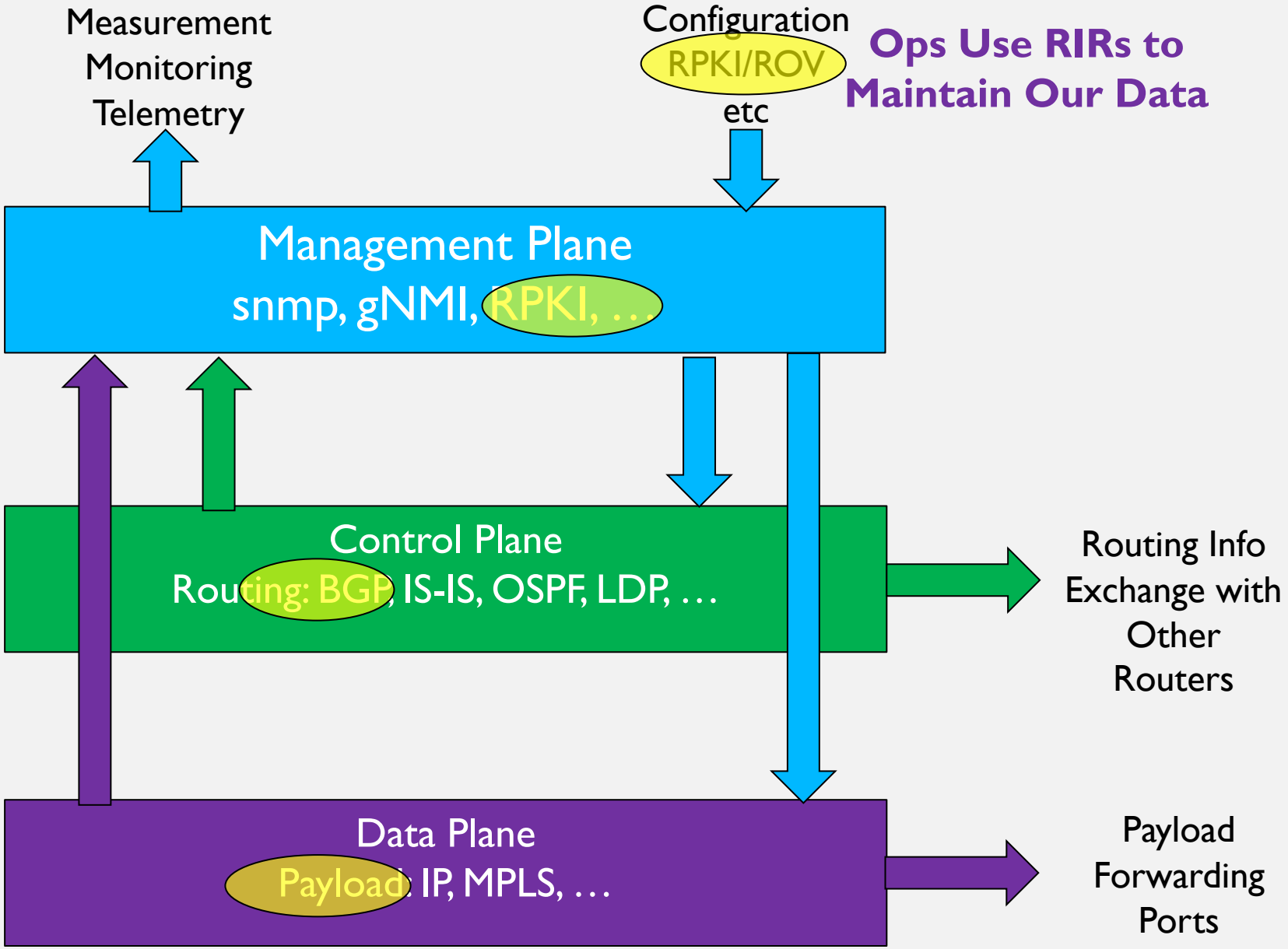
- No built-in validation that updates are legitimate
- The chain of trust spans continents
- Lack of reliable resource data

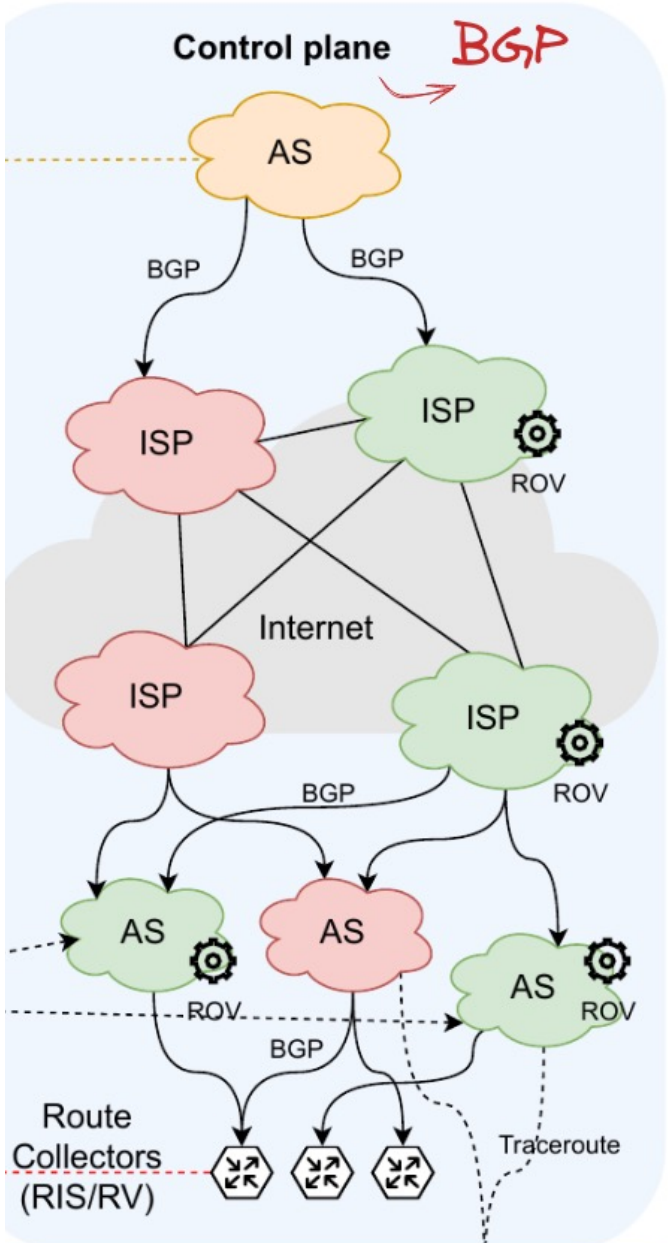
RPKI

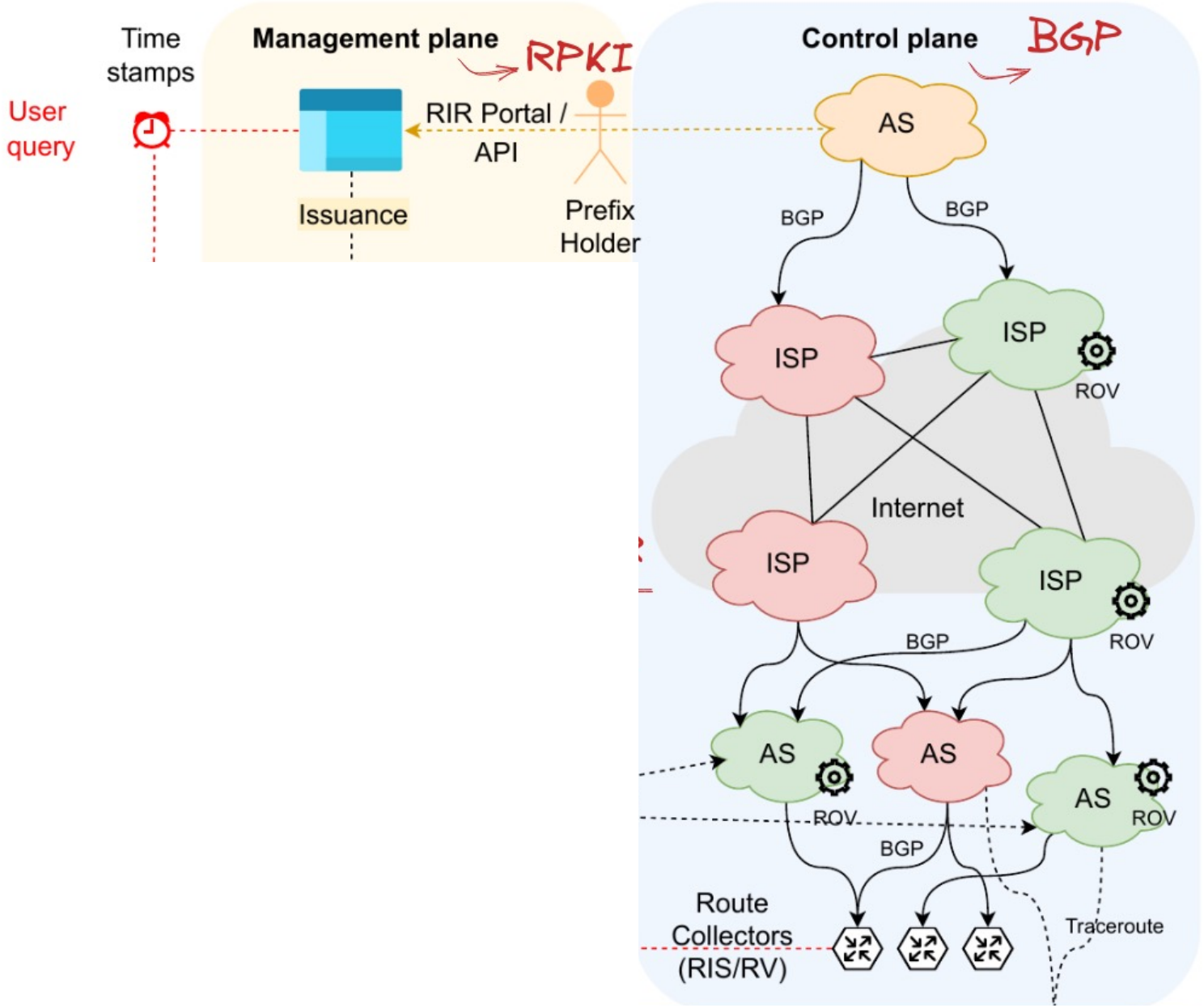
- RPKI + <as_path_security> - Our best chance to secure BGP
- RPKI is a distributed public key database to manage digitally signed objects
- ROA: Route Origin Authorization asserts which ASes are allowed to announce which prefixes
- ROV: Routers use ROA data to filter route (prefix, origin AS) → **not found**, **valid**, **invalid**

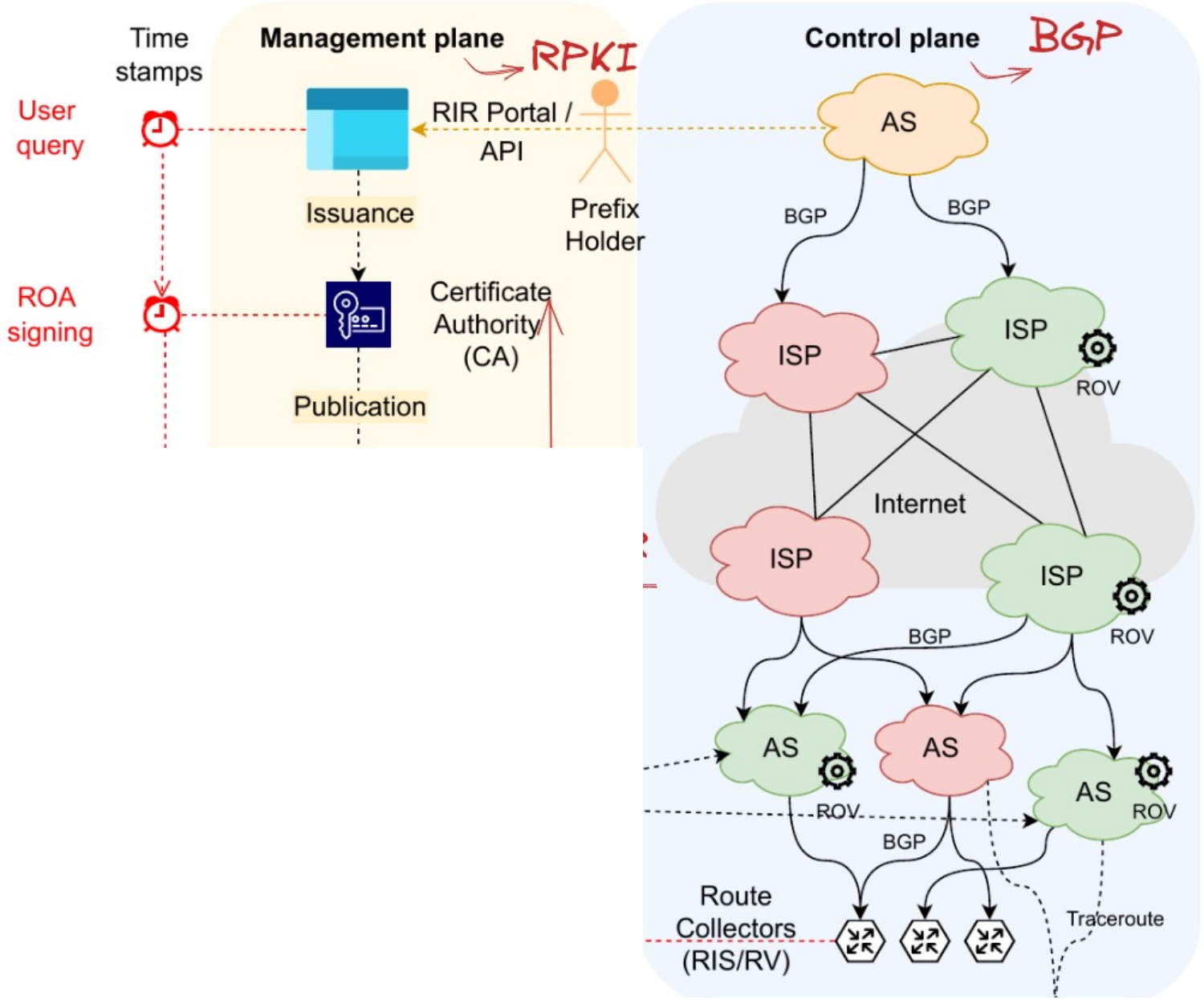


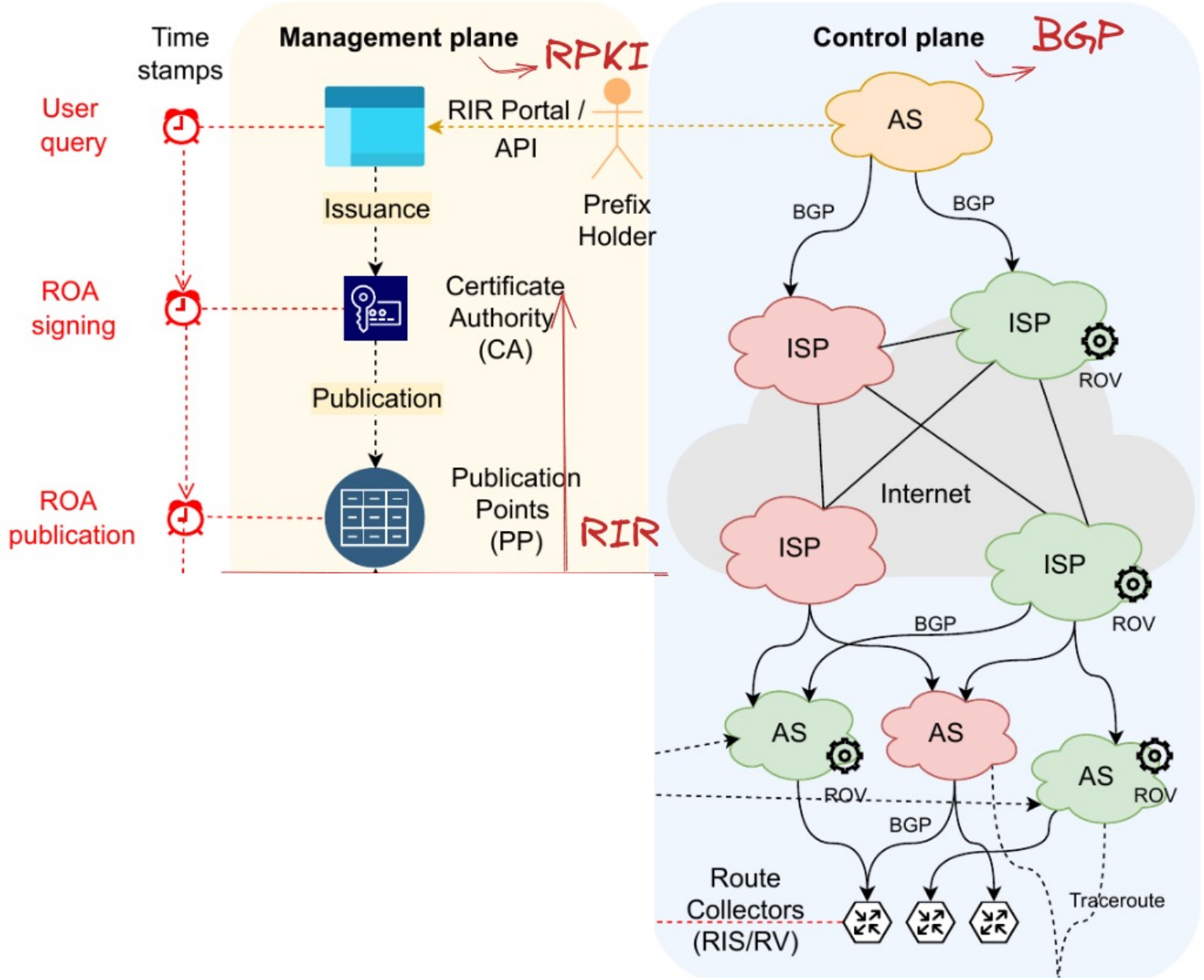


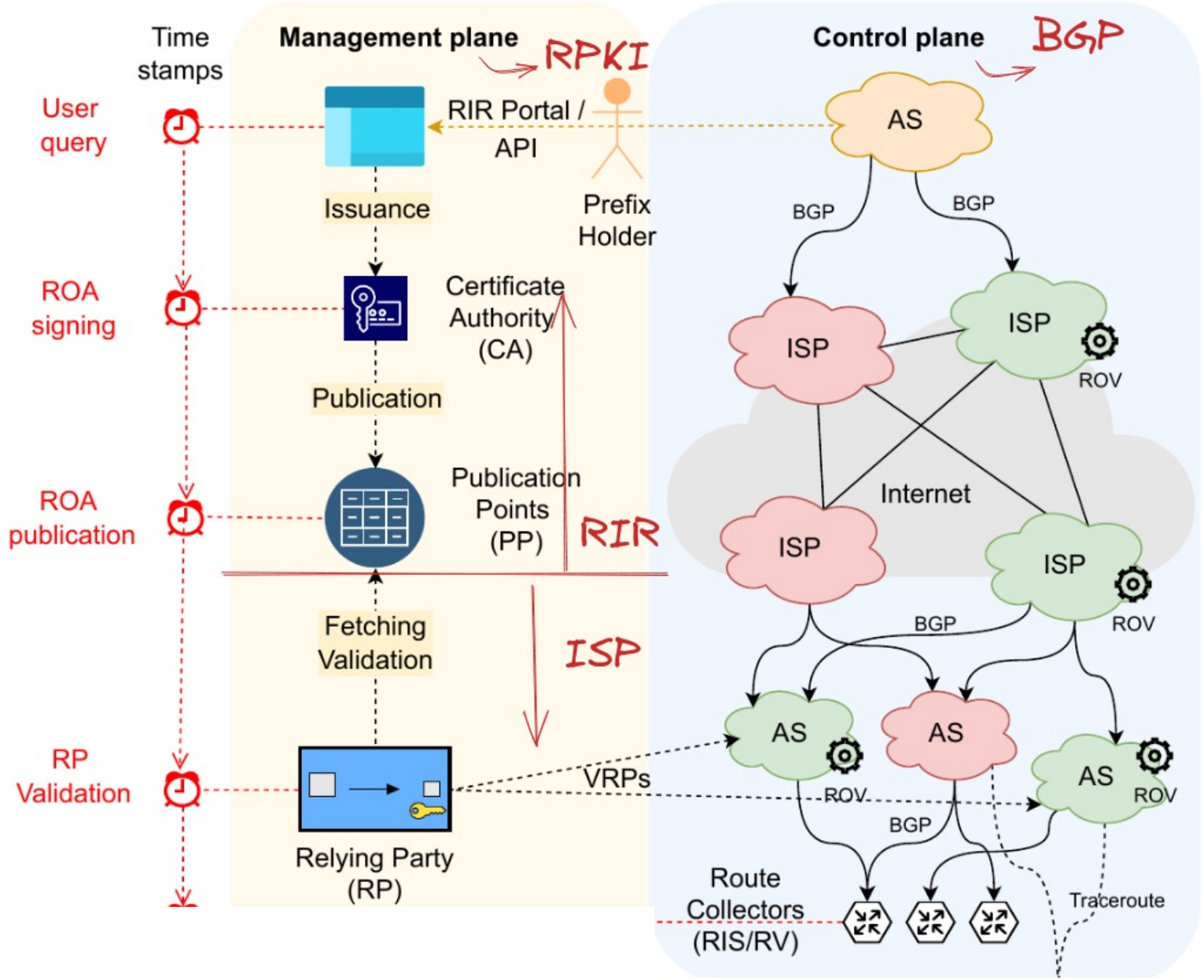


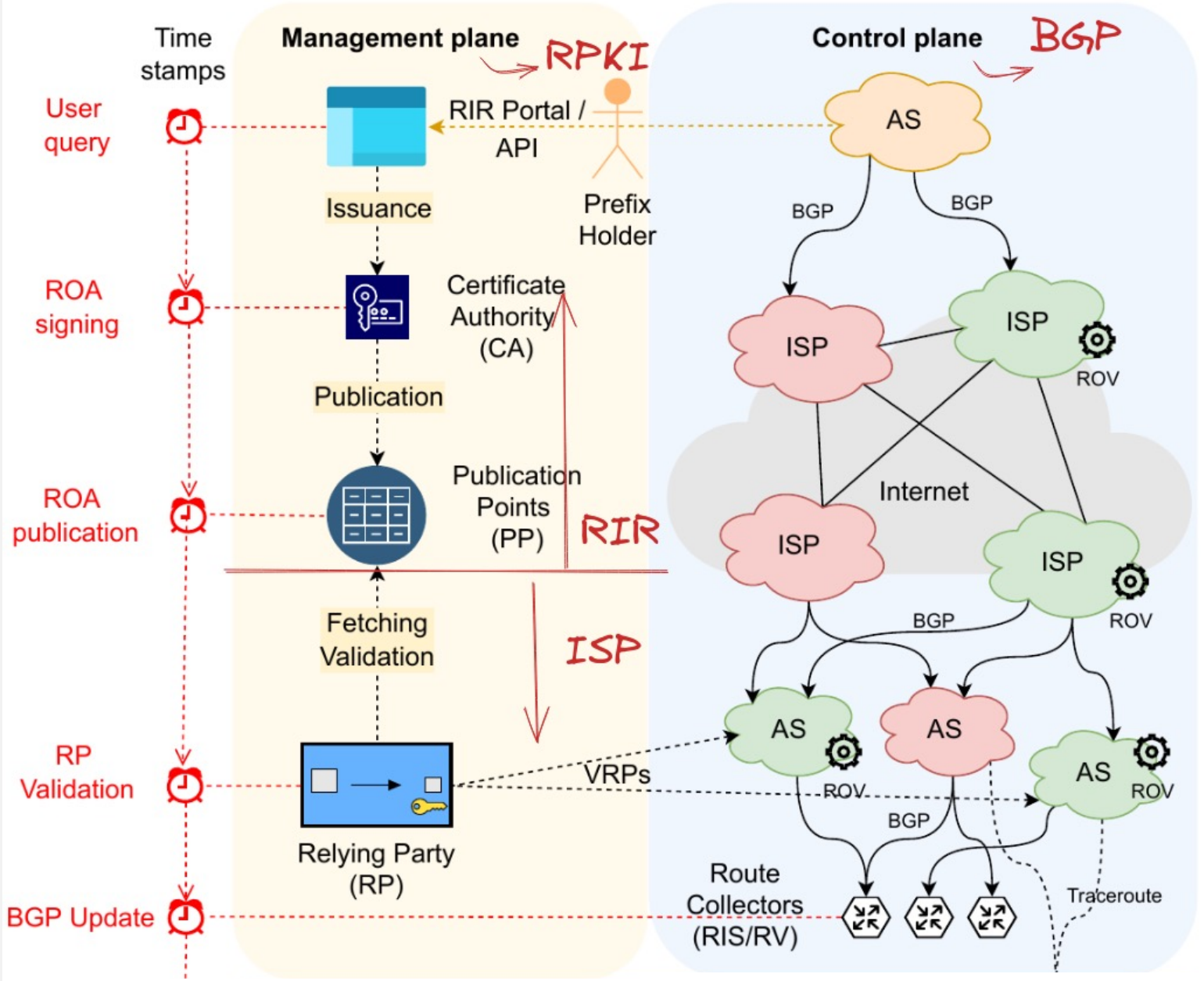

Control plane
BGP
AS
BGP
BGP
ISP
ISP
ROV
Internet
ISP
ISP
ROV
BGP
AS
AS
AS
ROV
ROV
BGP
Route
Collectors
(RIS/RV)
Traceroute

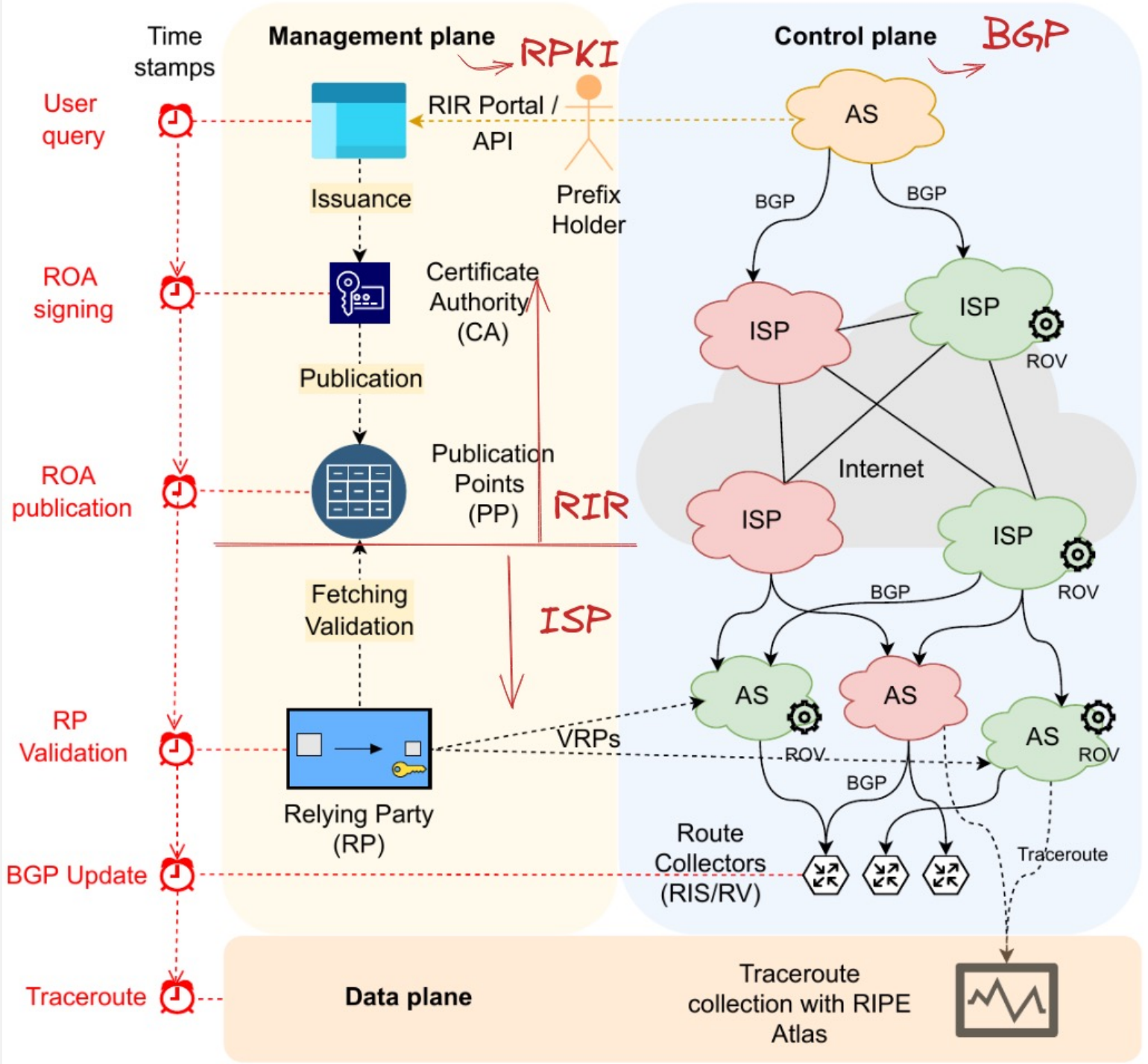

Time stamps
Management plane
RPKI
Control plane
BGP
User query
RIR Portal / API
AS
Issuance
Prefix Holder
BGP
BGP
ISP
ISP
ROV
Internet
ISP
ISP
ROV
BGP
AS
AS
AS
ROV
ROV
BGP
Route Collectors (RIS/RV)
Traceroute


Time stamps
Management plane
RPKI
Control plane
BGP
User query
RIR Portal / API
AS
Prefix Holder
BGP
BGP
Issuance
ISP
ISP
ROV
ROA signing
Certificate Authority (CA)
Internet
Publication
ISP
ISP
ROV
BGP
AS
AS
AS
ROV
ROV
BGP
Route Collectors (RIS/RV)
Traceroute


Time stamps
Management plane
RPKI
Control plane
BGP
User query
RIR Portal / API
AS
Issuance
Prefix Holder
BGP
BGP
ROA signing
Certificate Authority (CA)
ISP
ISP
ROV
Publication
Internet
ROA publication
Publication Points (PP)
RIR
ISP
ISP
ROV
BGP
AS
AS
AS
ROV
ROV
BGP
Route Collectors (RIS/RV)
Traceroute


Time stamps
Management plane
RPKI
Control plane
BGP
User query
RIR Portal / API
AS
Issuance
Prefix Holder
BGP
BGP
ROA signing
Certificate Authority (CA)
ISP
ISP
ROV
Publication
Internet
ROA publication
Publication Points (PP)
RIR
ISP
ISP
ROV
BGP
Fetching Validation
ISP
AS
AS
AS
ROV
ROV
RP Validation
VRPs
BGP
Relying Party (RP)
Route Collectors (RIS/RV)
Traceroute


Time stamps
Management plane
RPKI
Control plane
BGP
User query
RIR Portal / API
AS
Issuance
Prefix Holder
BGP
BGP
ROA signing
Certificate Authority (CA)
ISP
ISP
ROV
Publication
Internet
ROA publication
Publication Points (PP)
RIR
ISP
ISP
ROV
BGP
Fetching Validation
ISP
AS
AS
AS
ROV
ROV
RP Validation
VRPs
BGP
Relying Party (RP)
Route Collectors (RIS/RV)
Traceroute
BGP Update


Time stamps
Management plane
RPKI
Control plane
BGP
User query
RIR Portal / API
AS
Issuance
Prefix Holder
BGP
BGP
ROA signing
Certificate Authority (CA)
ISP
ISP
ROV
Publication
ROA publication
Publication Points (PP)
RIR
Internet
ISP
ISP
ROV
BGP
Fetching Validation
ISP
AS
AS
AS
ROV
ROV
RP Validation
VRPs
BGP
Relying Party (RP)
Route Collectors (RIS/RV)
Traceroute
BGP Update
Traceroute
Data plane
Traceroute collection with RIPE Atlas

THE EXPERIMENT(S)

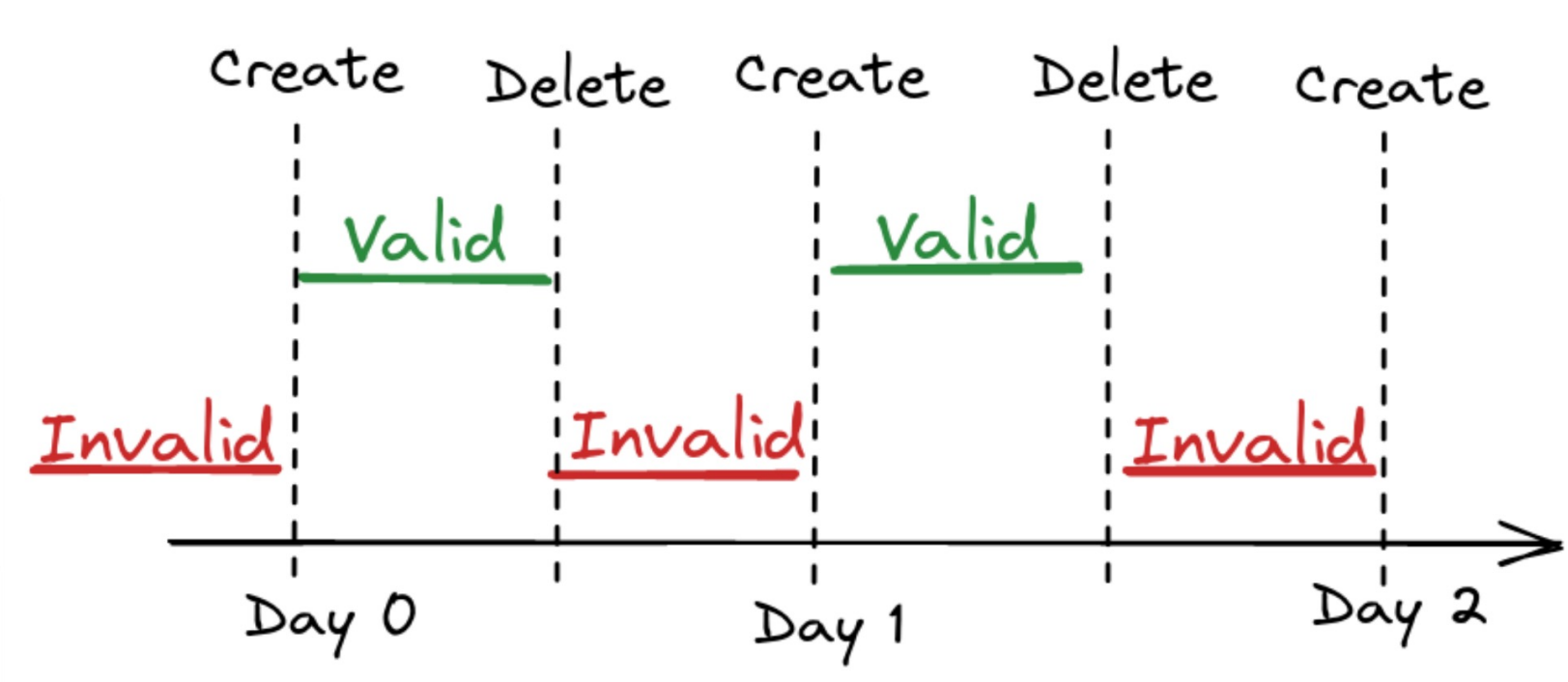
PREFIXES

- Each of the Five RIRs loaned us a few IPv4 /24s and IPv6 /48s
- Prefixes were announced from one AS with ROV upstreams and some direct IX peers which were non-ROV
- Measurements taken over almost a year

ROA BEACONS

- Used API or GUI at each RIR to Create and Delete ROAs
- Control /24s and /48s have non-varying 'good' ROAs, always Valid
- Test /24 and /48 always have an Invalidating ROA
- But Announced a Validating ROA once per day for half a day

ROA BEACONS



Invalidating ROA
AS😈
Prefix P

Validating ROA
AS😇
Prefix P

MEASUREMENT RELYING PARTY

- One instance of RP software
- See Philip Smith's measurements on how RPs vary ☹
- Did not run RPKI-RTR, because we were more interested in effect on BGP
- Some RPs have different implementation (single or multi-processing)

RIPE/RIS COLLECTORS

- Recorded Control and Test at RIPE/RIS
- If Control missing, that measurement is discarded
- This measures control plane, BGP, effect
- Used two collectors, RRC00 and RRC01. Studies have shown that's enough
- Has all the biases discussed for years

ROA CREATION DELAY

- Creation times vary significantly across RIRs, with medians ranging from a few minutes to over an hour for new ROAs to reach the publication points
- And we know of at least one NIR (not RIR) that only publishes once per day!

ROA CREATION DELAY (MIN)

	Sign*	NotBefore*	Publication†	Relying Party†	BGP‡
AFRINIC	0 (0)	0 (0)	3 (2)	14 (13)	15 (16)
APNIC	10 (13)	10 (13)	14 (16)	34 (38)	26 (28)
ARIN	- (-)	- (-)	69 (97)	81 (109)	95 (143)
LACNIC	0 (0)	- (-)	54 (32)	66 (42)	51 (34)
RIPE	0 (0)	0 (0)	4 (4)	14 (13)	18 (18)
After fix:					
ARIN	- (-)	- (-)	8 (9)	21 (22)	28 (23)

RIR ISP

- ARIN and LacNIC were signing in GMT (HSM)
  - But publishing in Local Time
  - So, NotBefore appeared to be hours before publication
  - We reported, they hacked a work-around
- APNIC always waited for 20-minute batches

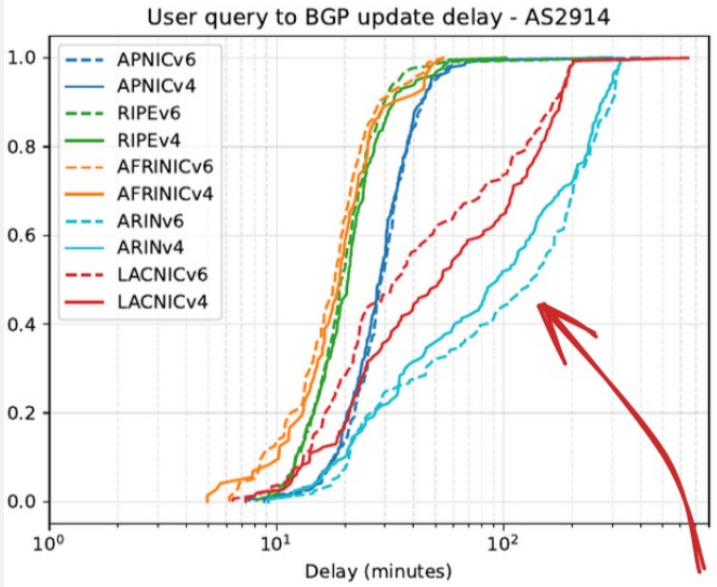
ROA REVOKE DELAY (MIN)

	Revocation*	Relying Party†	BGP‡
AFRINIC	0 (0)	13 (14)	34 (38)
APNIC	10 (12)	31 (36)	51 (56)
ARIN	0 (0)	14 (16)	45 (51)
LACNIC	0 (0)	18 (20)	48 (49)
RIPE	0 (0)	14 (13)	41 (50)

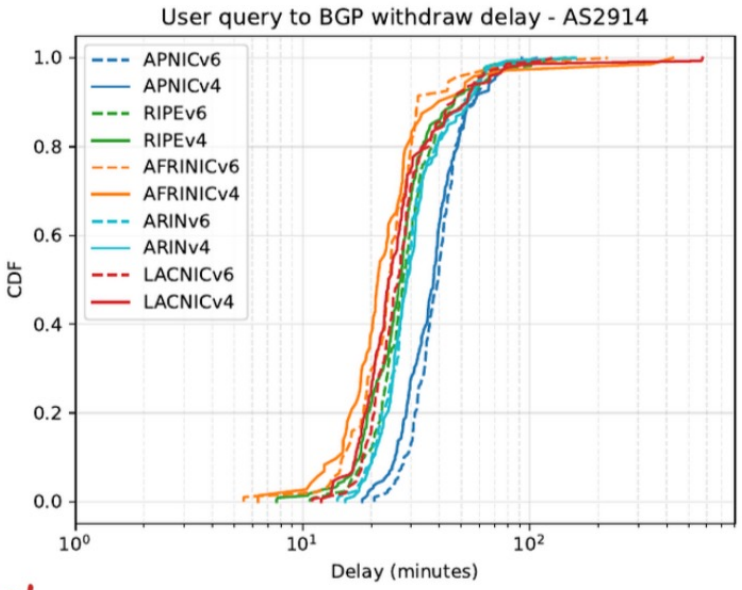
Additional APNIC delay possibly due to RP hanging Plus APNIC has that 20-minute batching delay

WITHDRAWS ARE SLOWER

- Because all of the router's / AS's RP caches must have received the Withdraw from the PPs
- ROV only needs one Validating ROA
- So only one cache needs to have a ROA for the router to Validate

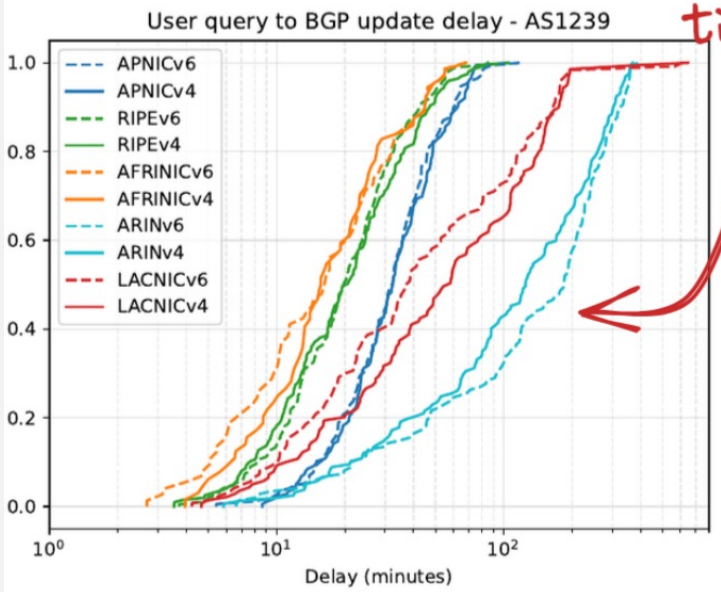


(a) ROA creation: NTT (AS2914).

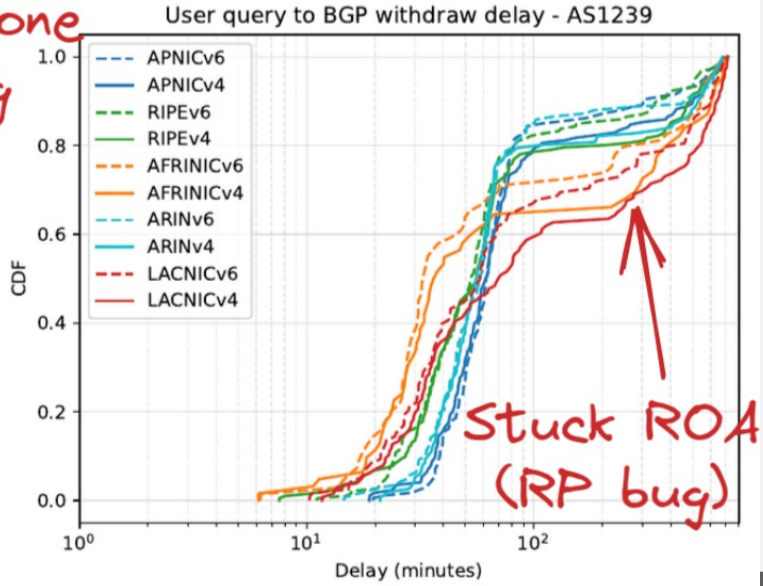


(b) ROA deletion: NTT (AS2914).

ARIN LACNIC timezone bug

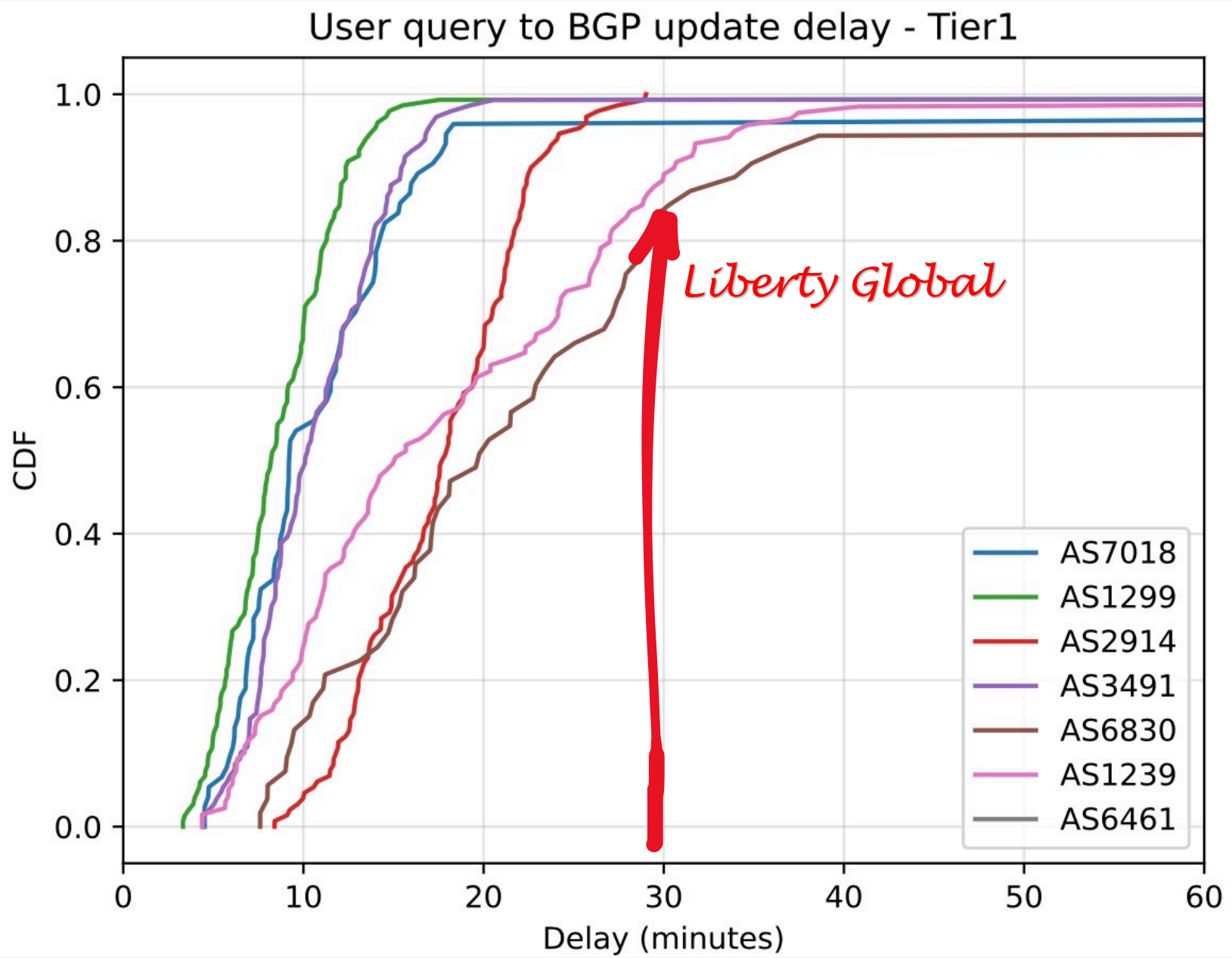


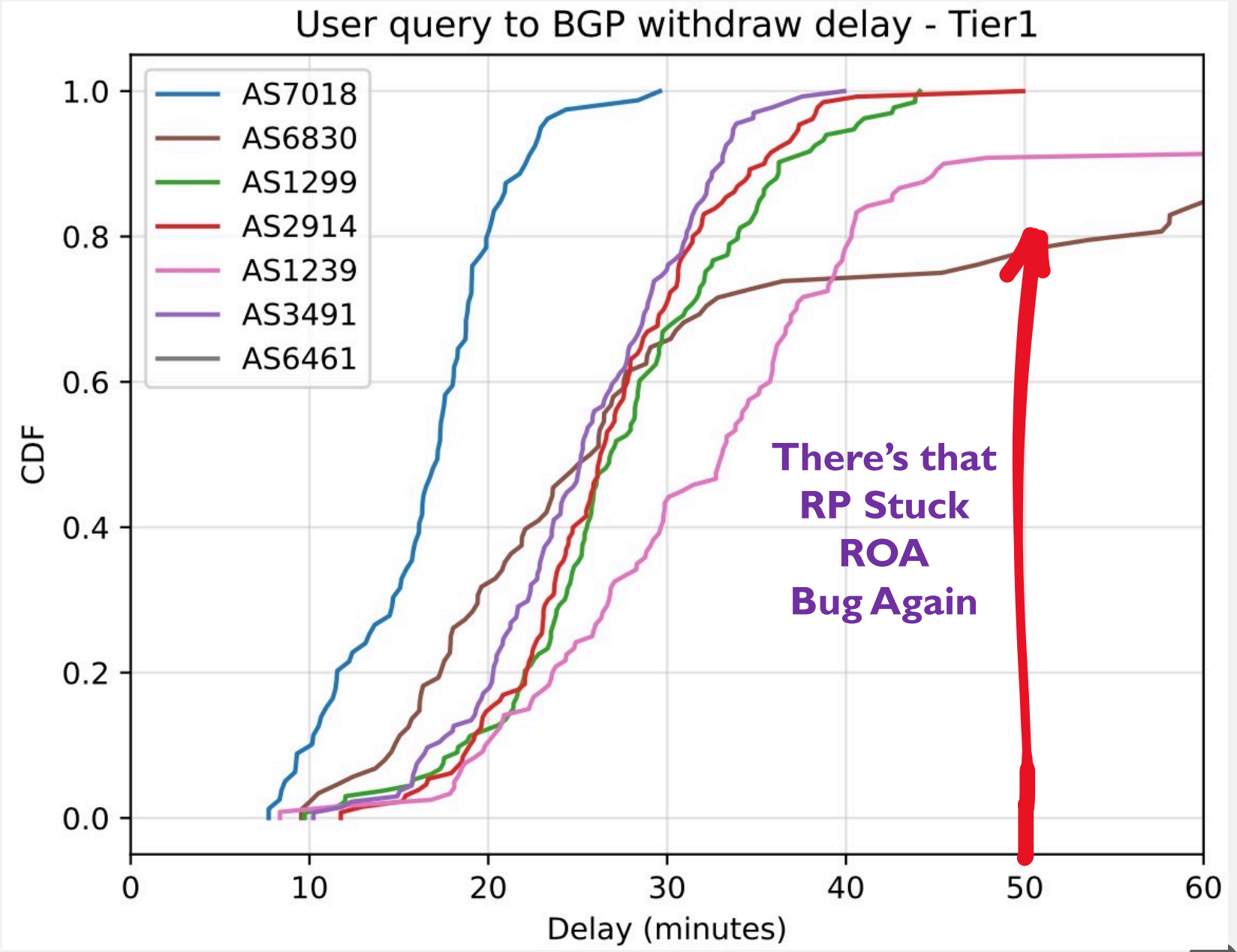
(c) ROA creation: Sprint (AS1239).



Stuck ROA (RP bug)

(d) ROA deletion: Sprint (AS1239).

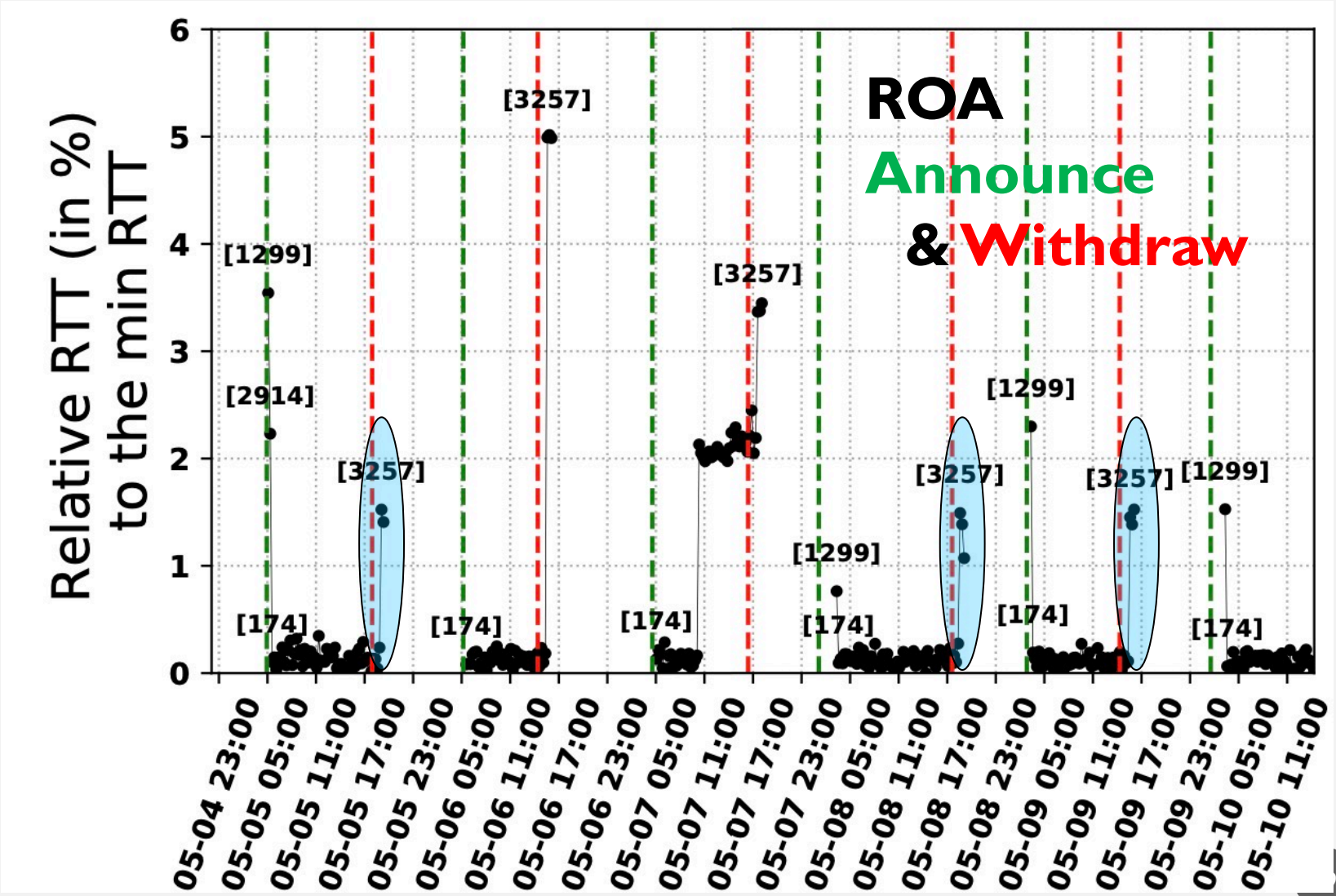




DATA PLANE MEASUREMENT

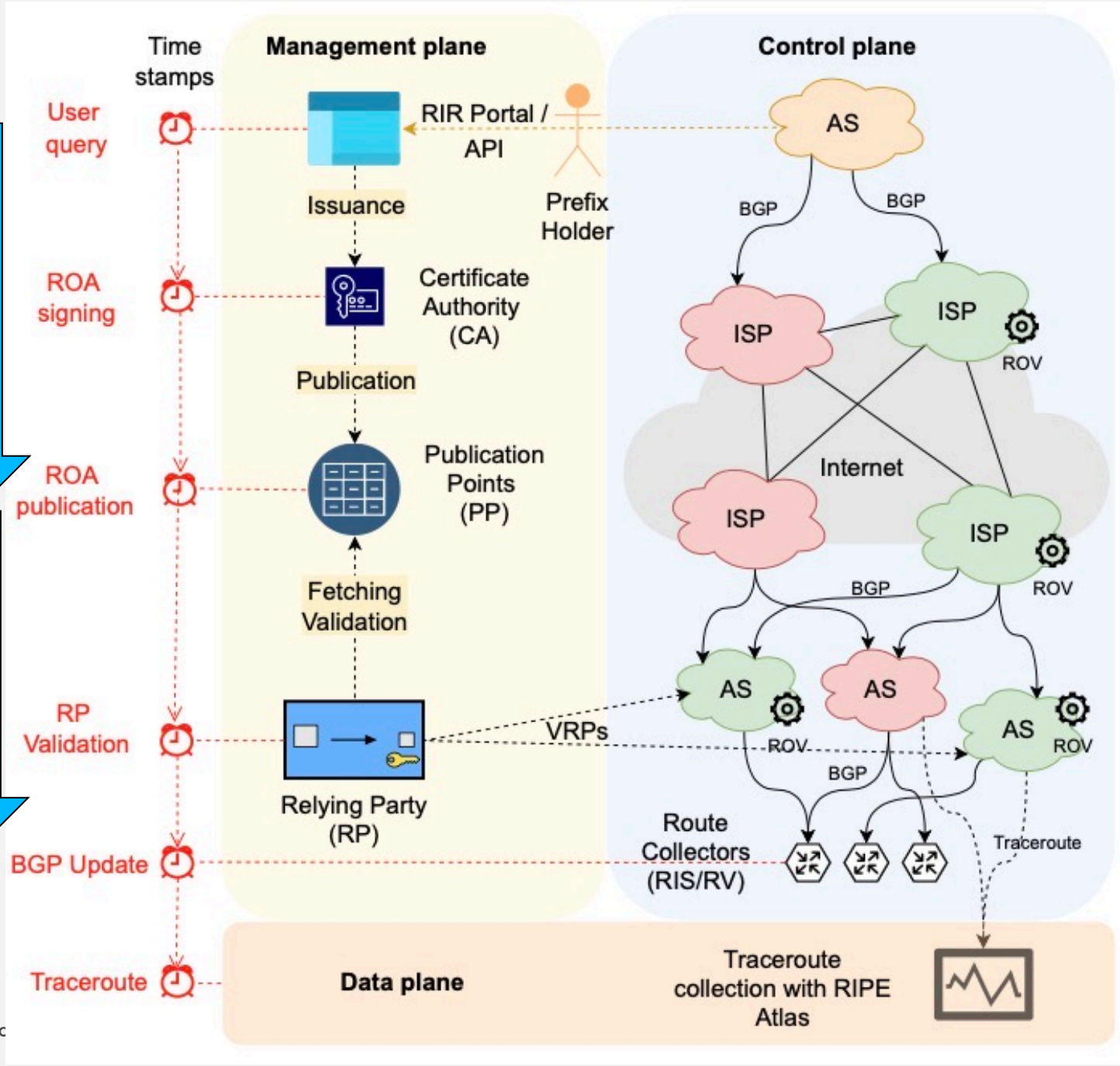
- Ran *traceroute* from Atlas Probes
- To the Test prefixes
- Every 15 minutes
- Result pretty much the same as BGP at RIPE/RIS, but
- Path hunting after a Withdraw is graphically obvious

DATA PLANE & PATH HUNTING



RIR Delay

ISP Delay



AND ISP DELAY LOOKS BIGGER

	Sign*	NotBefore*	Publication†	Relying Party†	BGP‡
AFRINIC	0 (0)	0 (0)	3 (2)	14 (13)	15 (16)
APNIC	10 (13)	10 (13)	14 (16)	34 (38)	26 (28)
ARIN	- (-)	- (-)	69 (97)	81 (109)	95 (143)
LACNIC	0 (0)	- (-)	54 (32)	66 (42)	51 (34)
RIPE	0 (0)	0 (0)	4 (4)	14 (13)	18 (18)
After fix:					
ARIN	- (-)	- (-)	8 (9)	21 (22)	28 (23)

Let's assume ARIN and LacNIC TimeZone anomalies are fixed

PROBLEMS

- BGP propagates in minutes. RPKI propagates in O(hour). This has business impacts, e.g.
  - Time to Repair for a bad ROA
  - Time to authorize a DDoS mitigator
- Two RIRs with HSM in GMT and CAs in Local Time Zone. Reported and 'fixed'
- Some RPs have not discovered fork() and exec()
- ROA Anatomy varies between RIRs

RPKI Time-of-Flight: Tracking Delays in the Management, Control, and Data Planes

Romain Fontugne[1], Amreesh Phokeer[2], Cristel Pelsser[3], Kevin Vermeulen[4], and Randy Bush[1,5]

[1] IIJ Research Lab romain@iij.ad.jp
[2] Internet Society phokeer@isoc.org
[3] UCLouvain cristel.pelsser@uclouvain.be
[4] LAAS-CNRS, Université de Toulouse, CNRS, Toulouse, France
kevin.vermeulen@laas.fr
[5] Arrcus, Inc randy@psg.com

Abstract. As RPKI is becoming part of ISPs' daily operations and Route Origin Validation is getting widely deployed, one wonders how long it takes for the effect of RPKI changes to appear in the data plane.

https://www.manrs.org/wp-content/uploads/2023/02/rpki-time-of-flight-pam23.pdf

QUESTIONS?