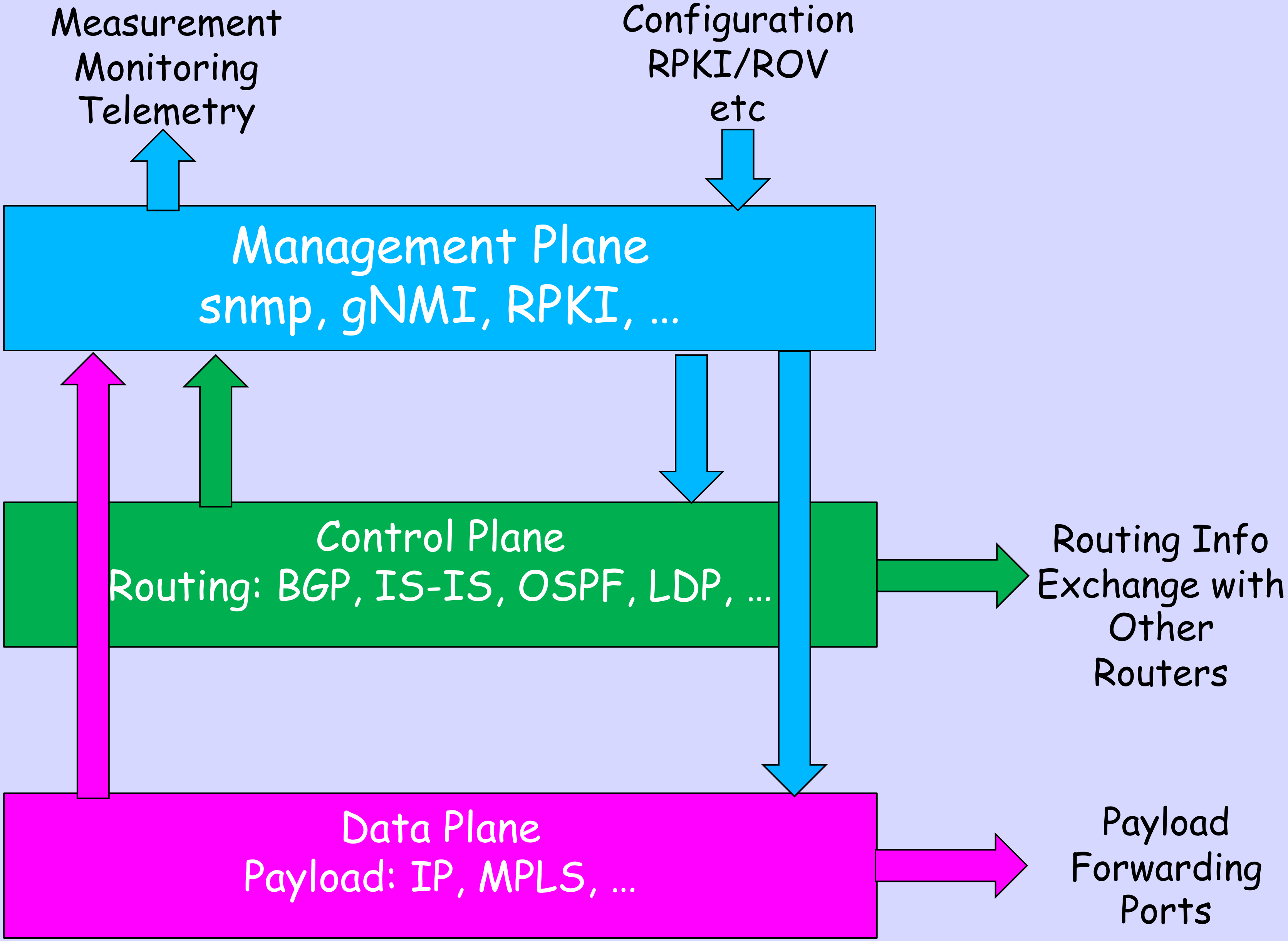
RPKI Ecosystem Measurement

IEPG

2023.03.26

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



Measurement
Monitoring
Telemetry

Configuration
RPKI/ROV
etc

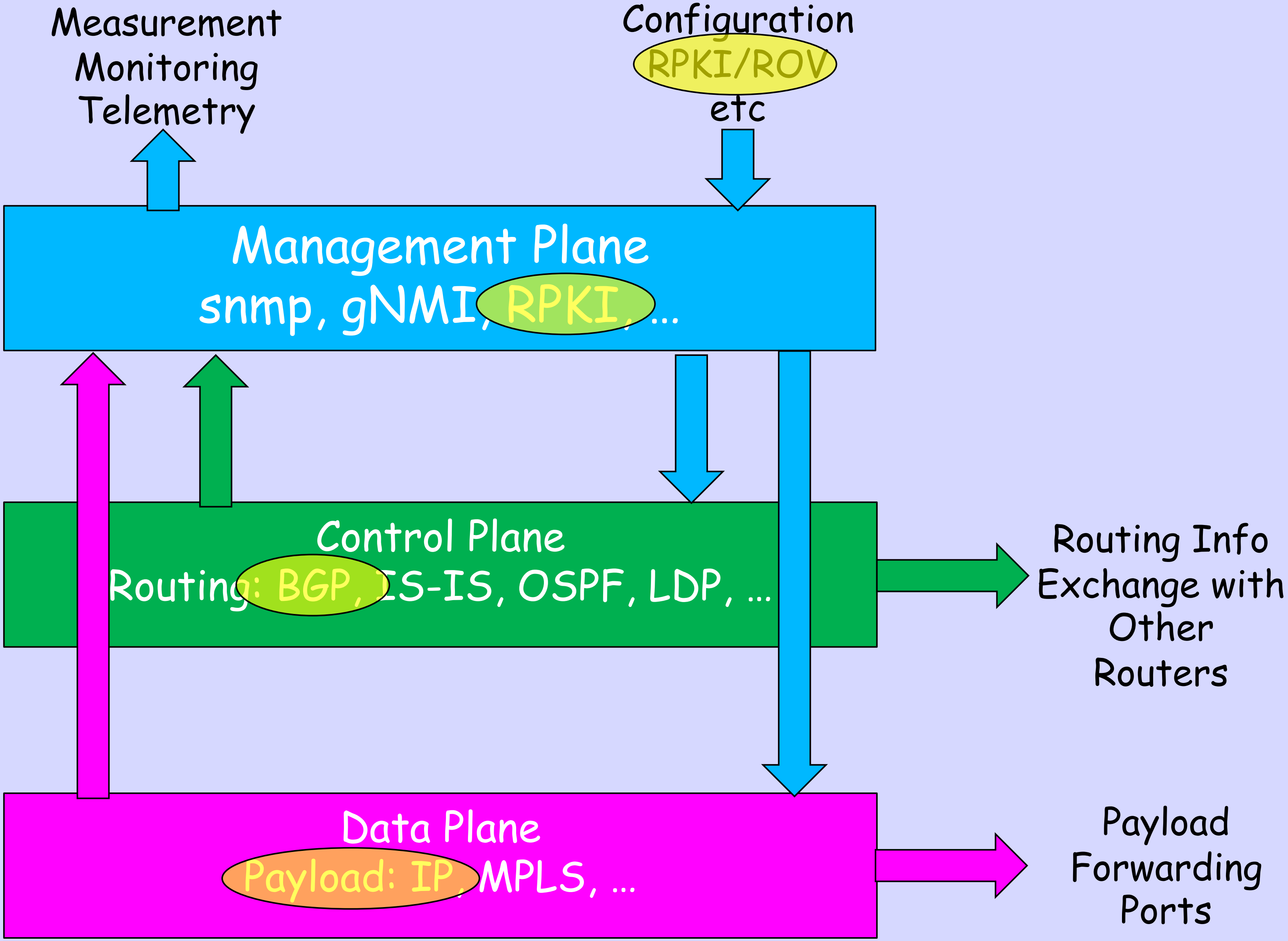
Management Plane
snmp, gNMI, RPKI, …

Control Plane
Routing: BGP, IS-IS, OSPF, LDP, …

Routing Info
Exchange with
Other
Routers

Data Plane
Payload: IP, MPLS, …

Payload
Forwarding
Ports



Measurement
Monitoring
Telemetry

Configuration
RPKI/ROV
etc

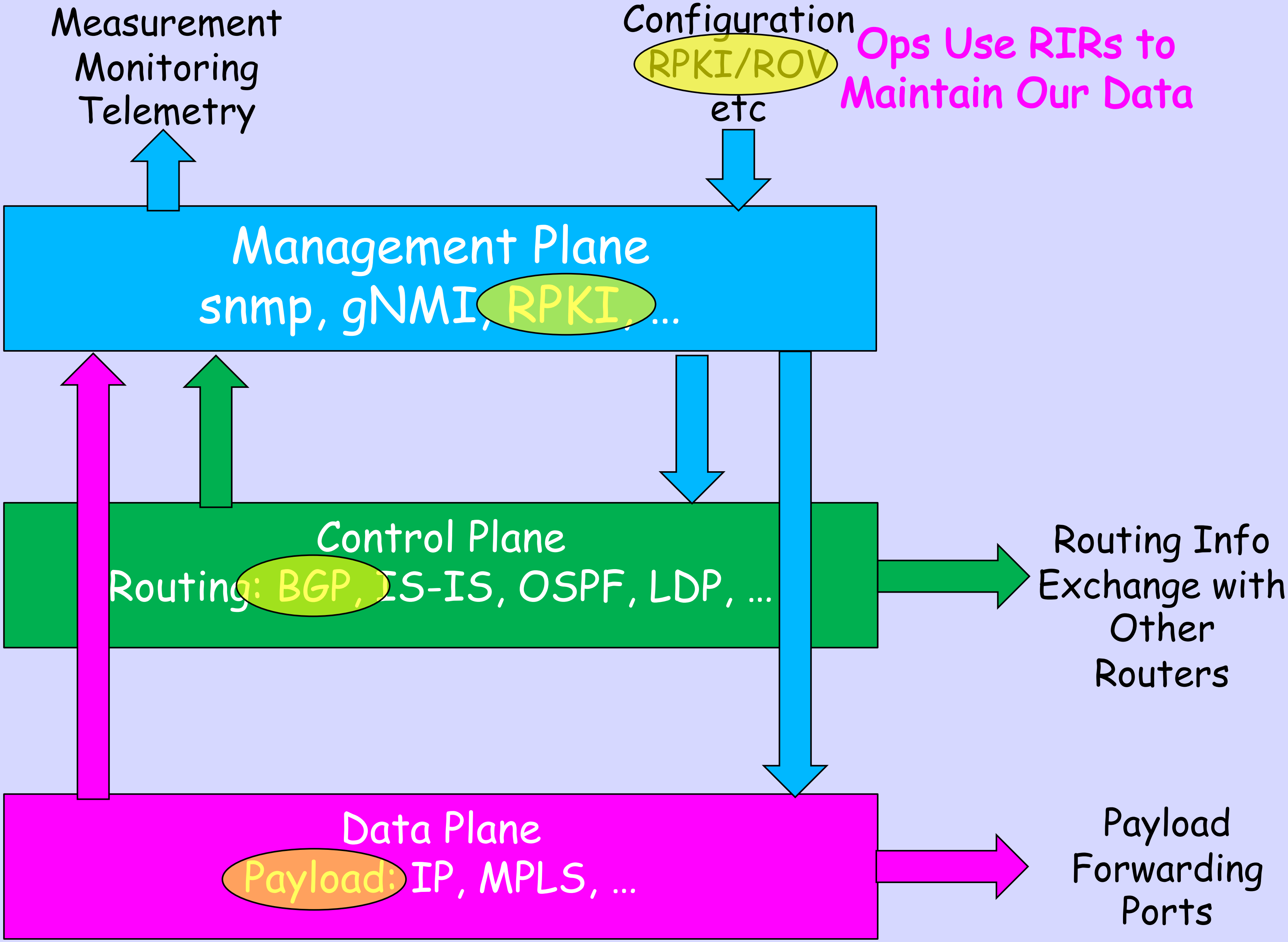
Management Plane
snmp, gNMI, RPKI, …

Control Plane
Routing: BGP, IS-IS, OSPF, LDP, …

Routing Info
Exchange with
Other
Routers

Data Plane
Payload: IP, MPLS, …

Payload
Forwarding
Ports



Measurement
Monitoring
Telemetry

Configuration
RPKI/ROV
etc

Ops Use RIRs to
Maintain Our Data

Management Plane
snmp, gNMI, RPKI, …

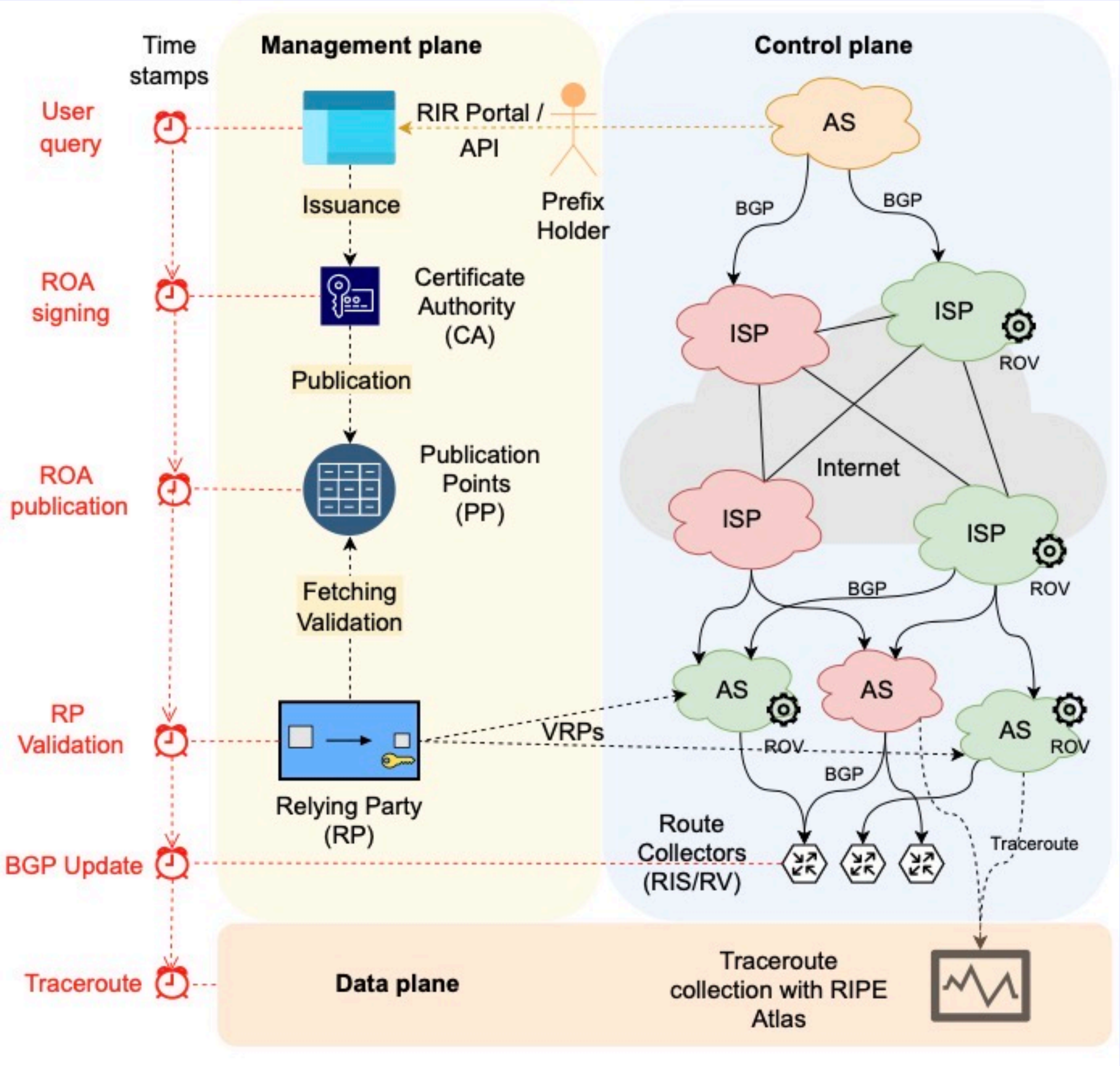
Control Plane
Routing: BGP, IS-IS, OSPF, LDP, …

Routing Info
Exchange with
Other
Routers

Data Plane
Payload: IP, MPLS, …

Payload
Forwarding
Ports

The Money is Here ☺


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

The Experiment(s)

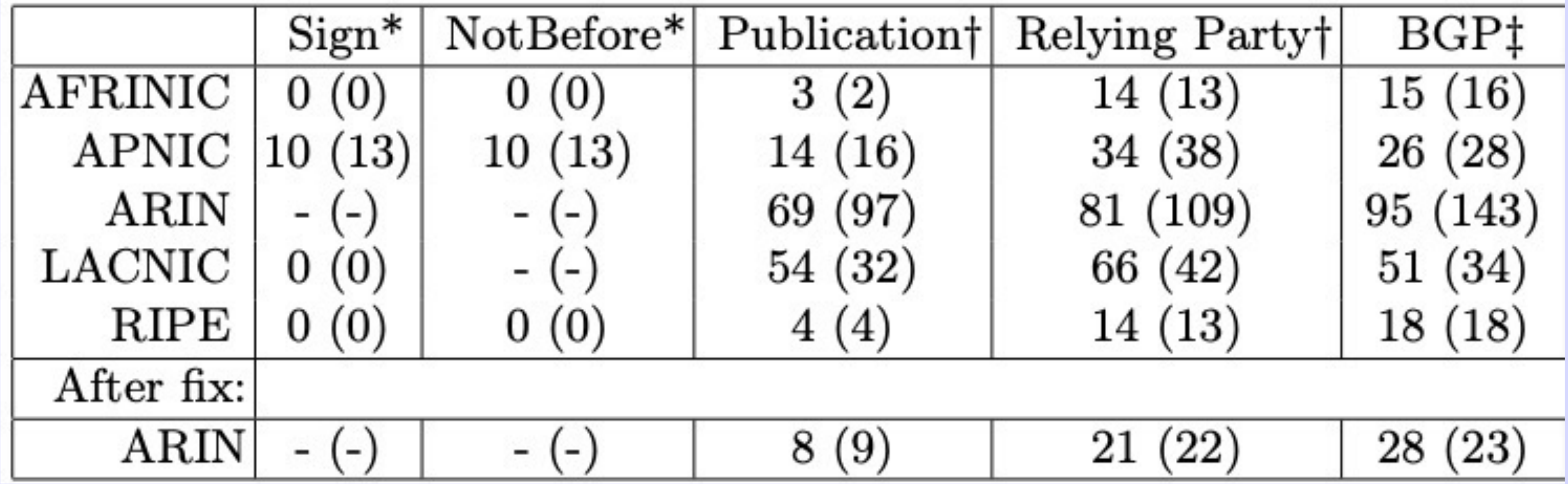
Prefixes

- 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
- Another set of RIPE prefixes from 3 ASs fed by non-ROV upstreams
- Measurements taken over eleven months

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 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)

- 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 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!

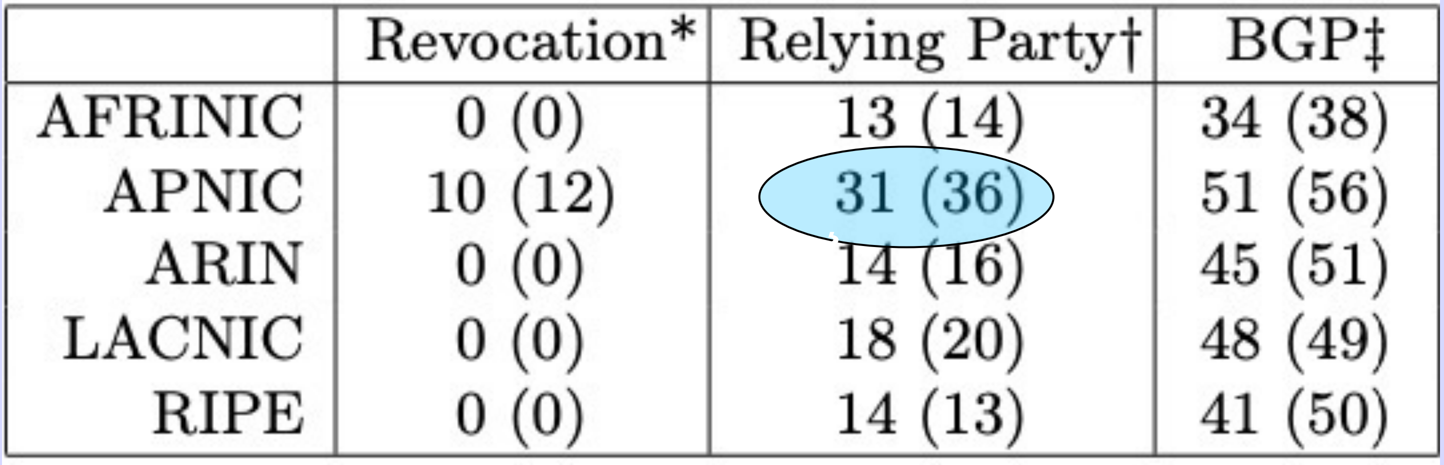
Measurement Relying Party

- 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 not discovered fork() and exec(), so HeadOfLineBlock trying to fetch from bad Publication Points

RIPE/RIS Collectors

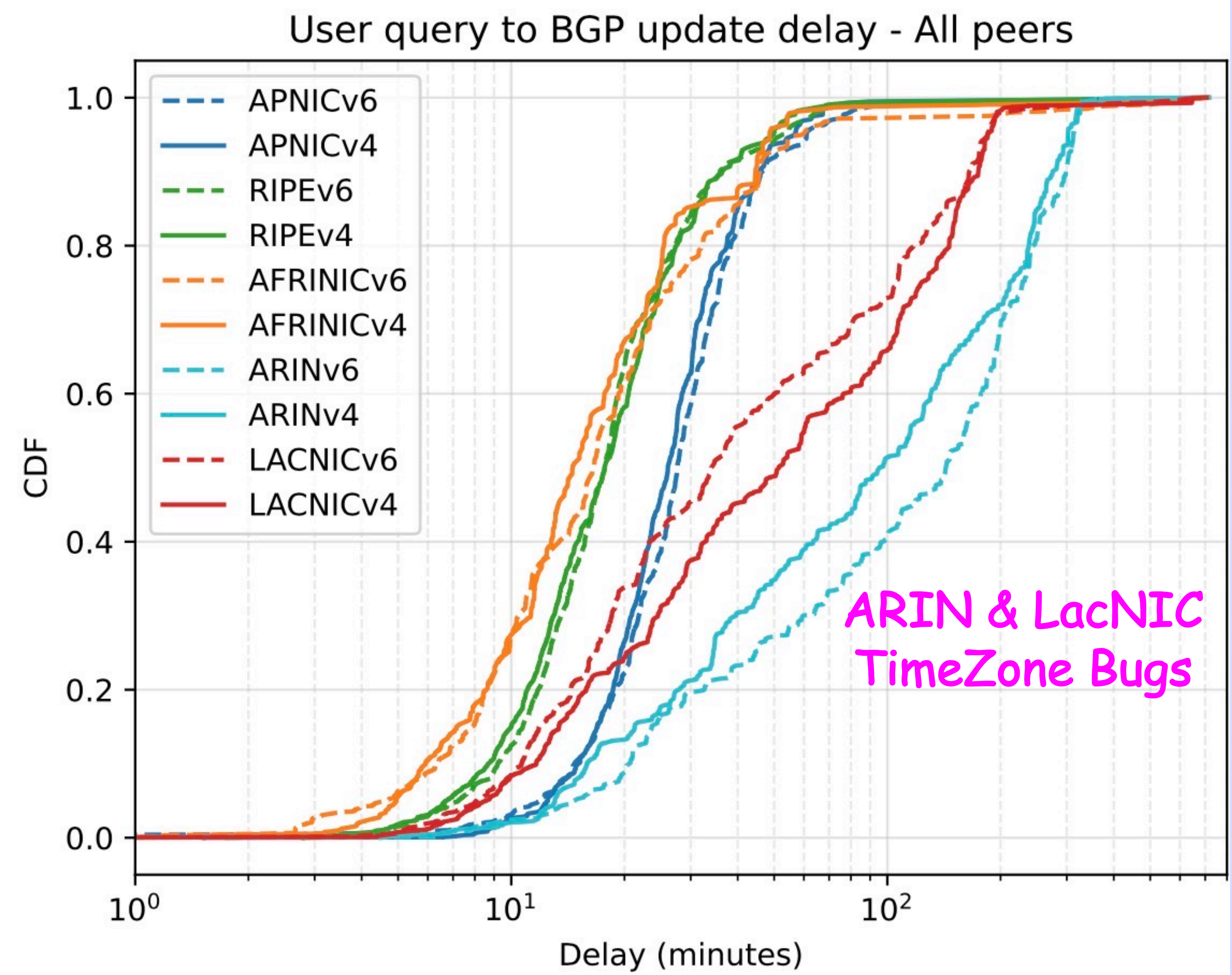
- Recorded Control and Test at RIPE/RIS
- If Control missing, that measurement is discarded
- This measures control plane, BGP, effect
- Used two collectors, RRC00 and RRC01. Studies have shown that's enough
- Has all the biases discussed for years

ROA Revoke Delay (min)

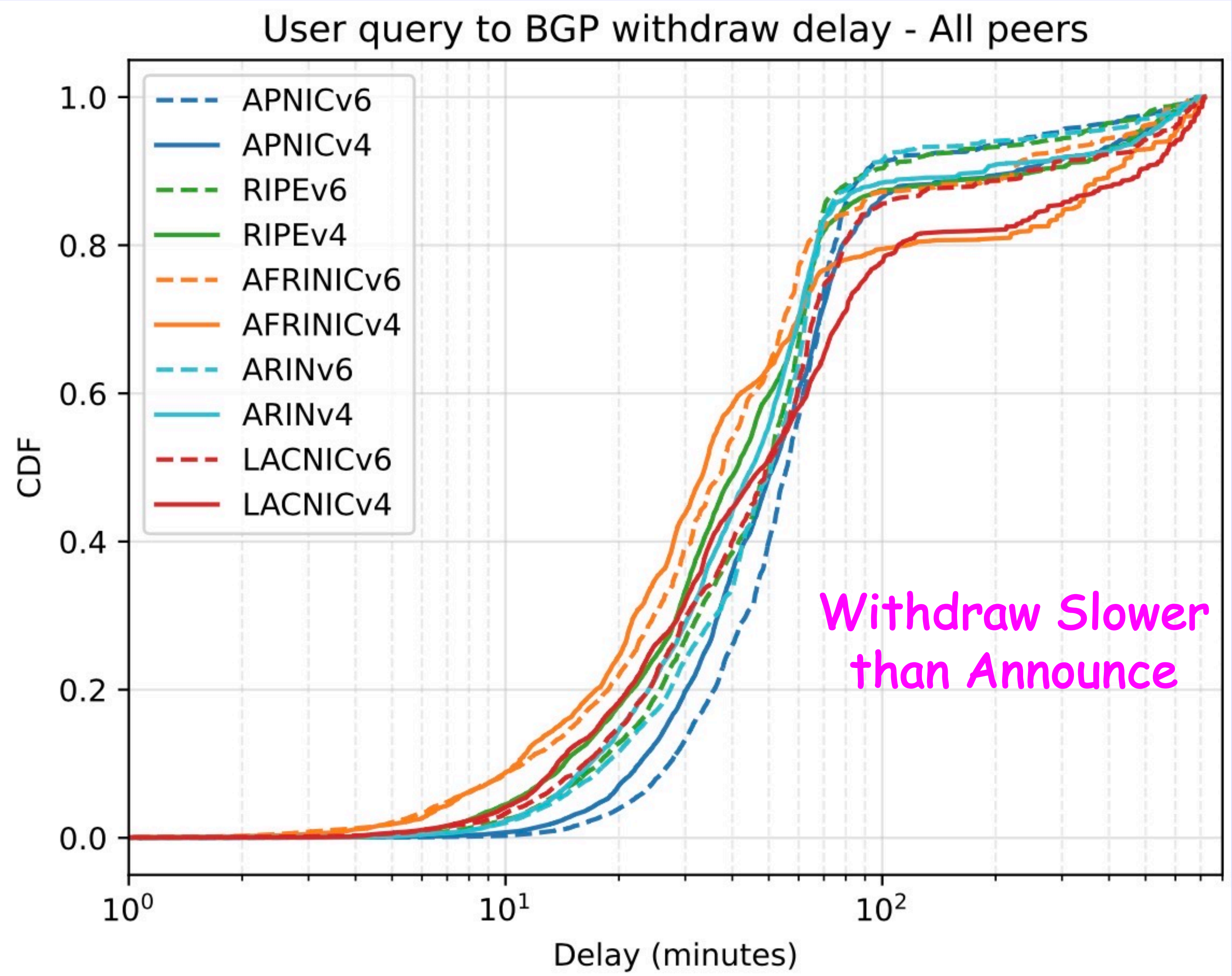


	Revocation*	Relying Party†	BGP‡
AFRINIC	0 (0)	13 (14)	34 (38)
APNIC	10 (12)	31 (36)	51 (56)
ARIN	0 (0)	14 (16)	45 (51)
LACNIC	0 (0)	18 (20)	48 (49)
RIPE	0 (0)	14 (13)	41 (50)

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



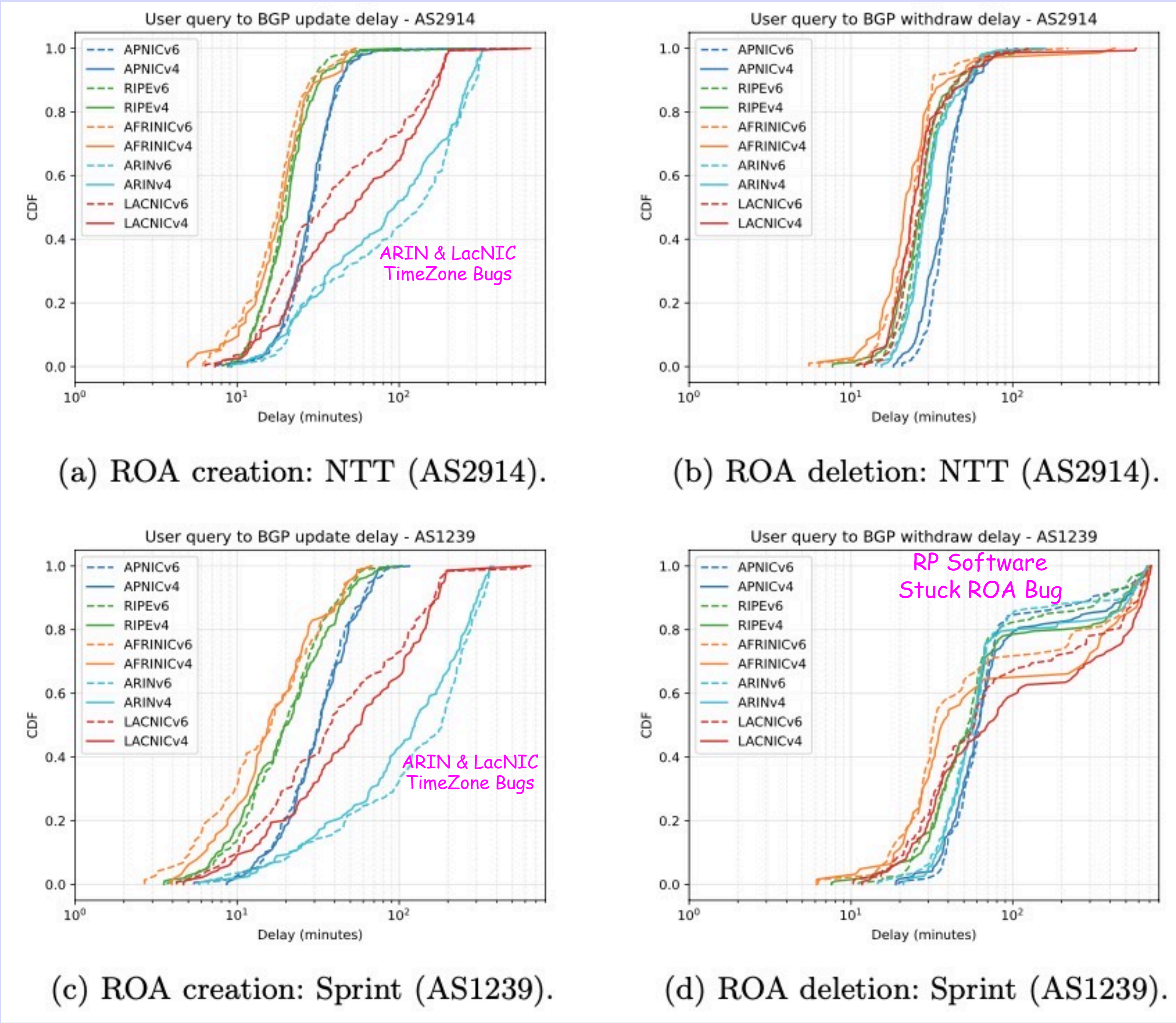
ARIN & LacNIC TimeZone Bugs



Withdraw Slower than Announce

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



ARIN & LacNIC TimeZone Bugs

(a) ROA creation: NTT (AS2914).

(b) ROA deletion: NTT (AS2914).

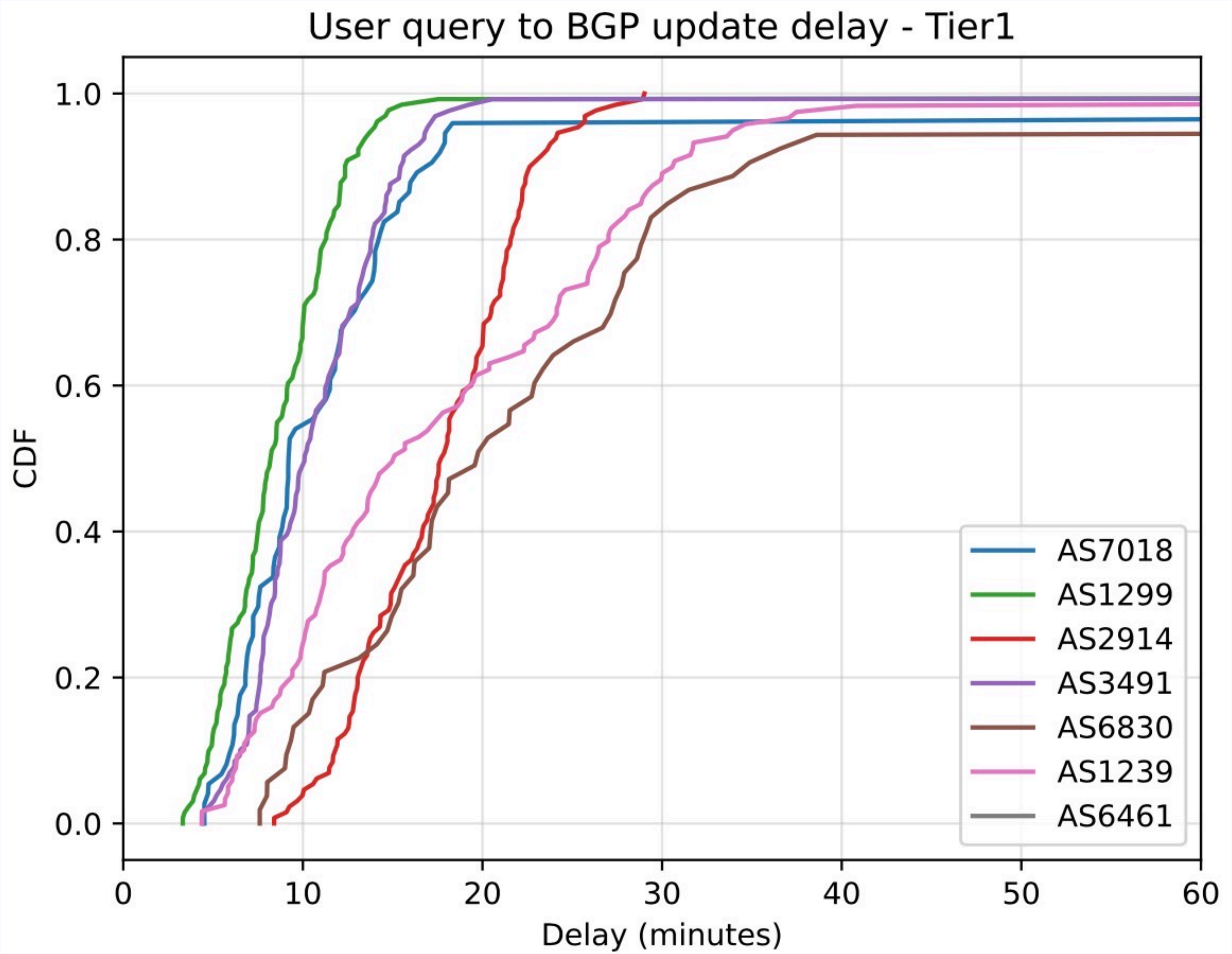
RP Software Stuck ROA Bug

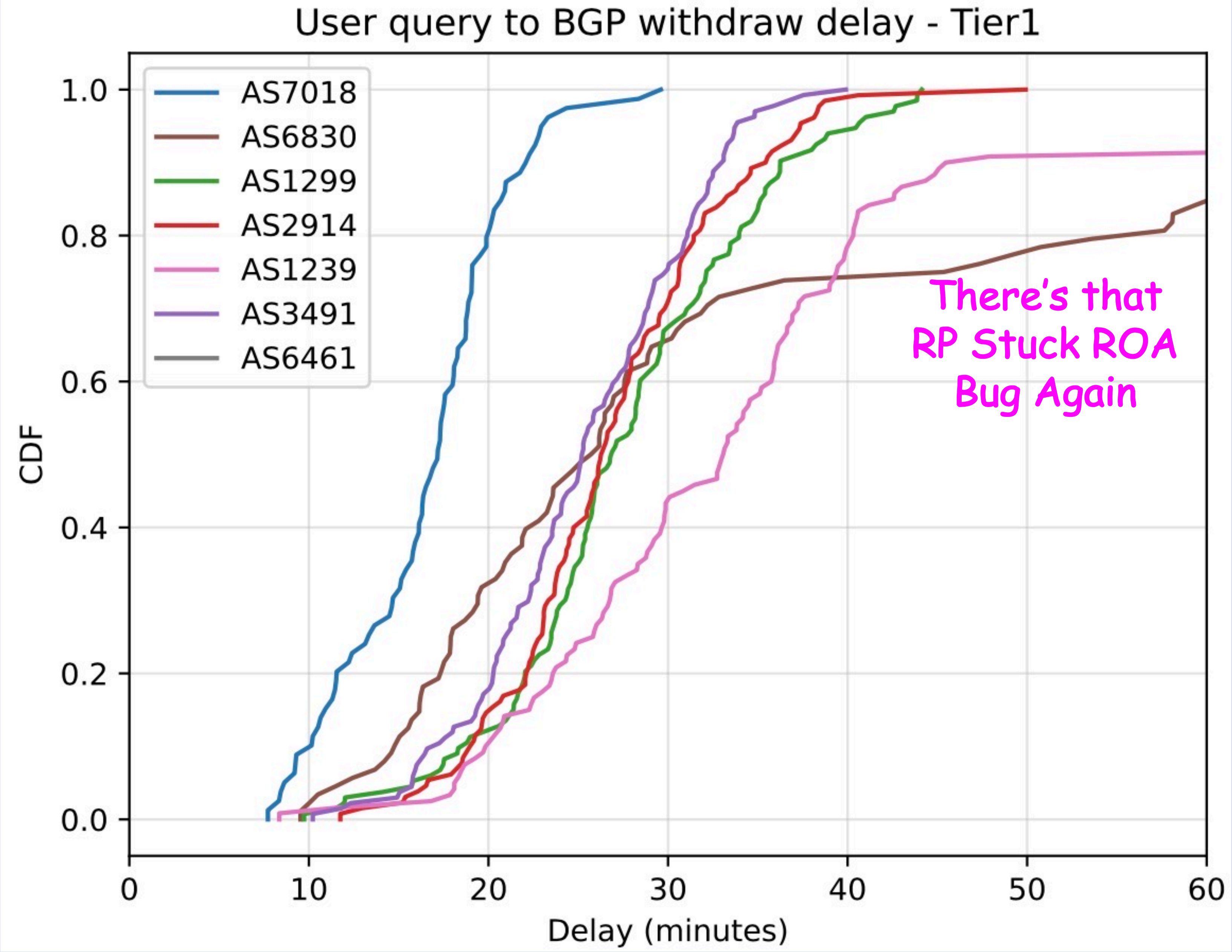
ARIN & LacNIC TimeZone Bugs

(c) ROA creation: Sprint (AS1239).

(d) ROA deletion: Sprint (AS1239).

Sprint's slower curve because RPs pull less frequently than NTT's, &/or sucky RP software
Sprint starts a bit earlier because routers poll RP caches more frequently than NTT's
Confirmed with Sprint and NTT


User query to BGP update delay - Tier1
CDF
0.0
0.2
0.4
0.6
0.8
1.0
Delay (minutes)
0
10
20
30
40
50
60
AS7018
AS1299
AS2914
AS3491
AS6830
AS1239
AS6461

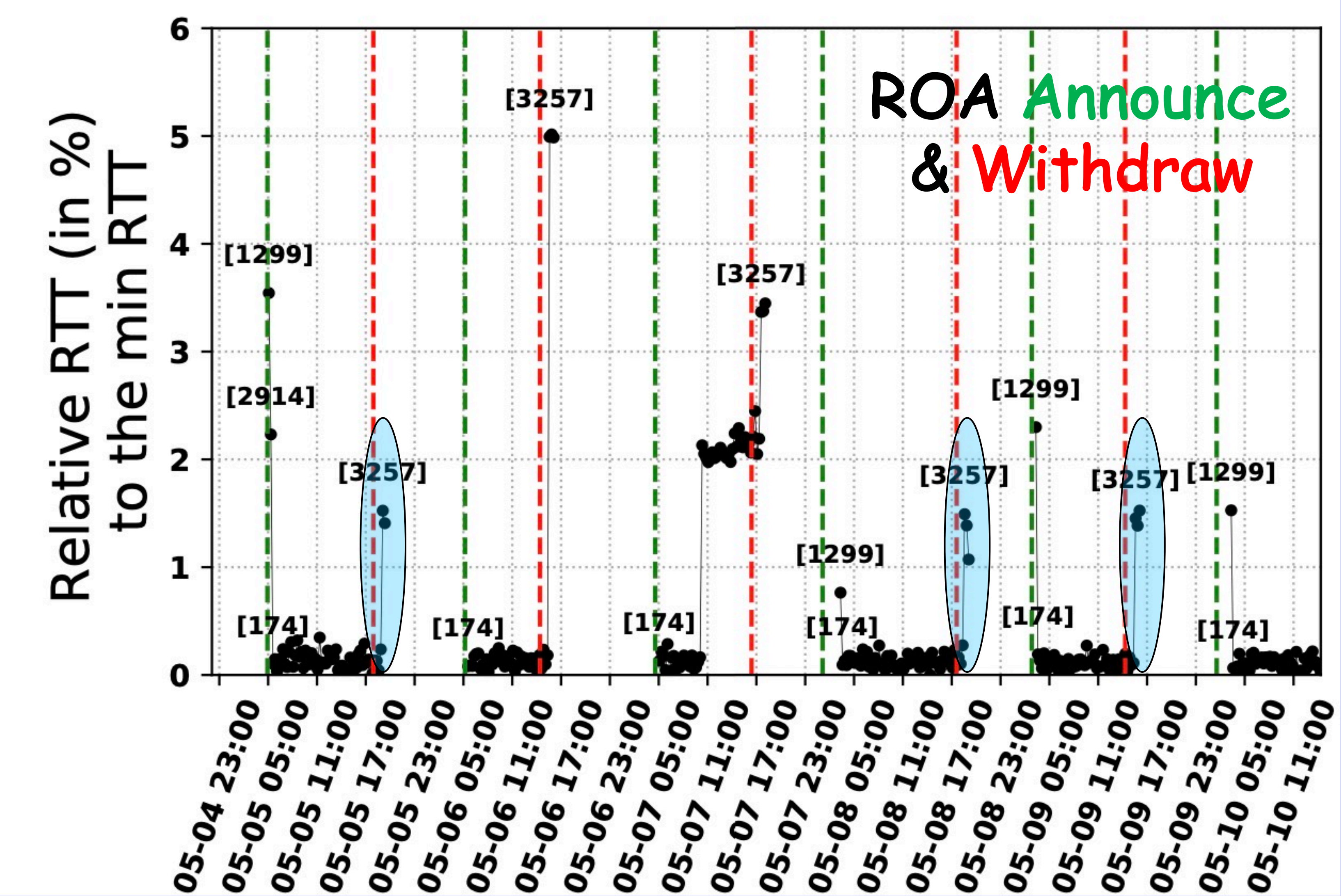


There's that RP Stuck ROA Bug Again

Data Plane Measurement

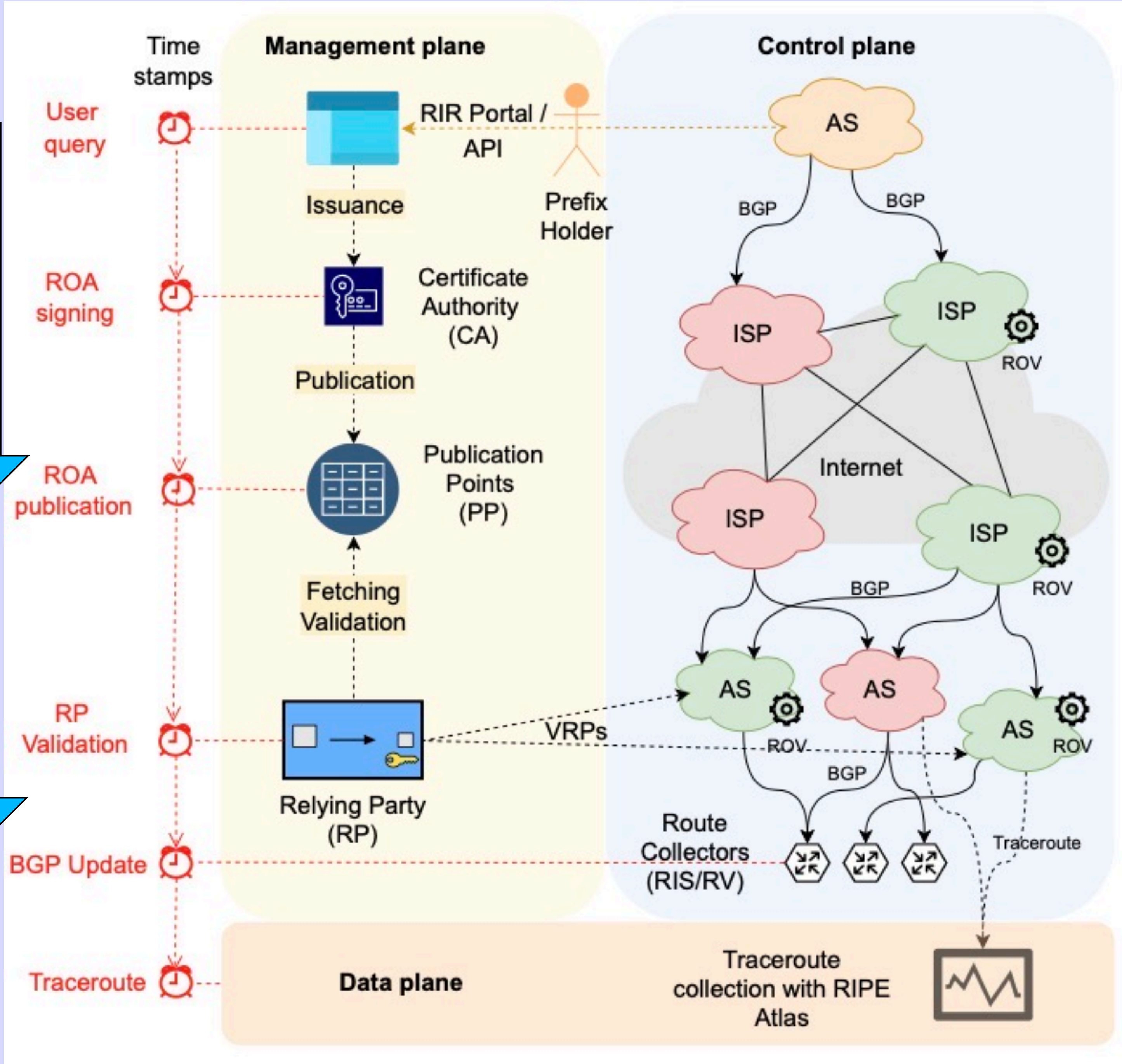
- Ran *traceroute* from Atlas Probes
- To the Test prefixes
- Every 15 minutes
- Result pretty much the same as BGP at RIPE/RIS, but
- Path hunting after a Withdraw is graphically obvious

Data Plane & Path Hunting

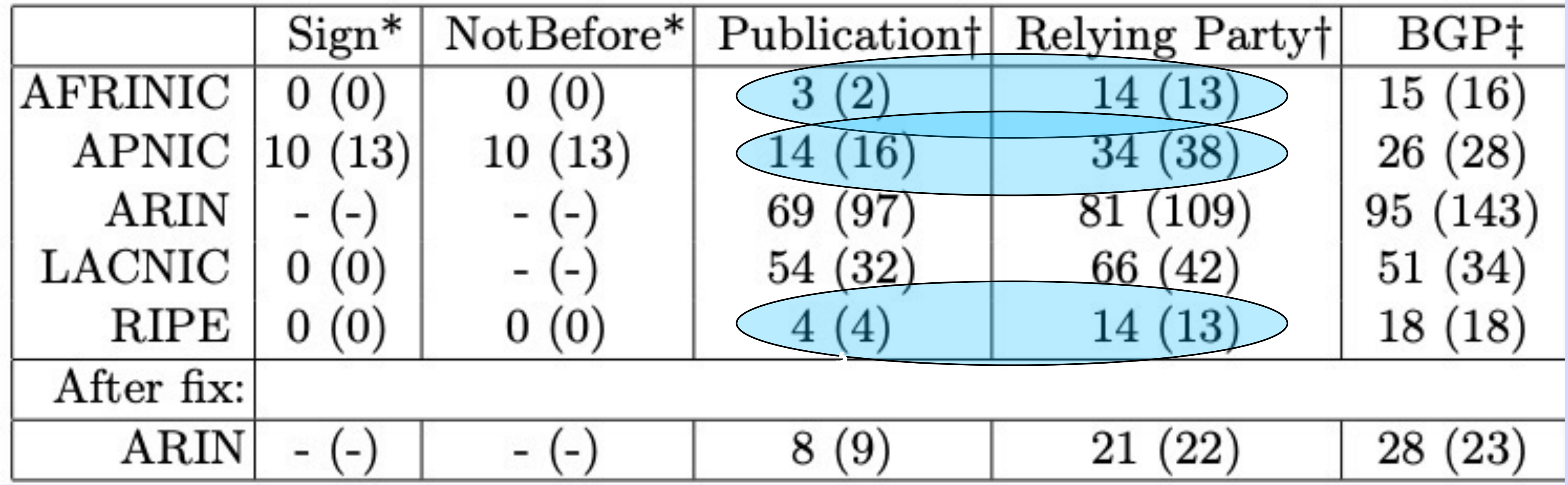


RIR Delay

ISP Delay



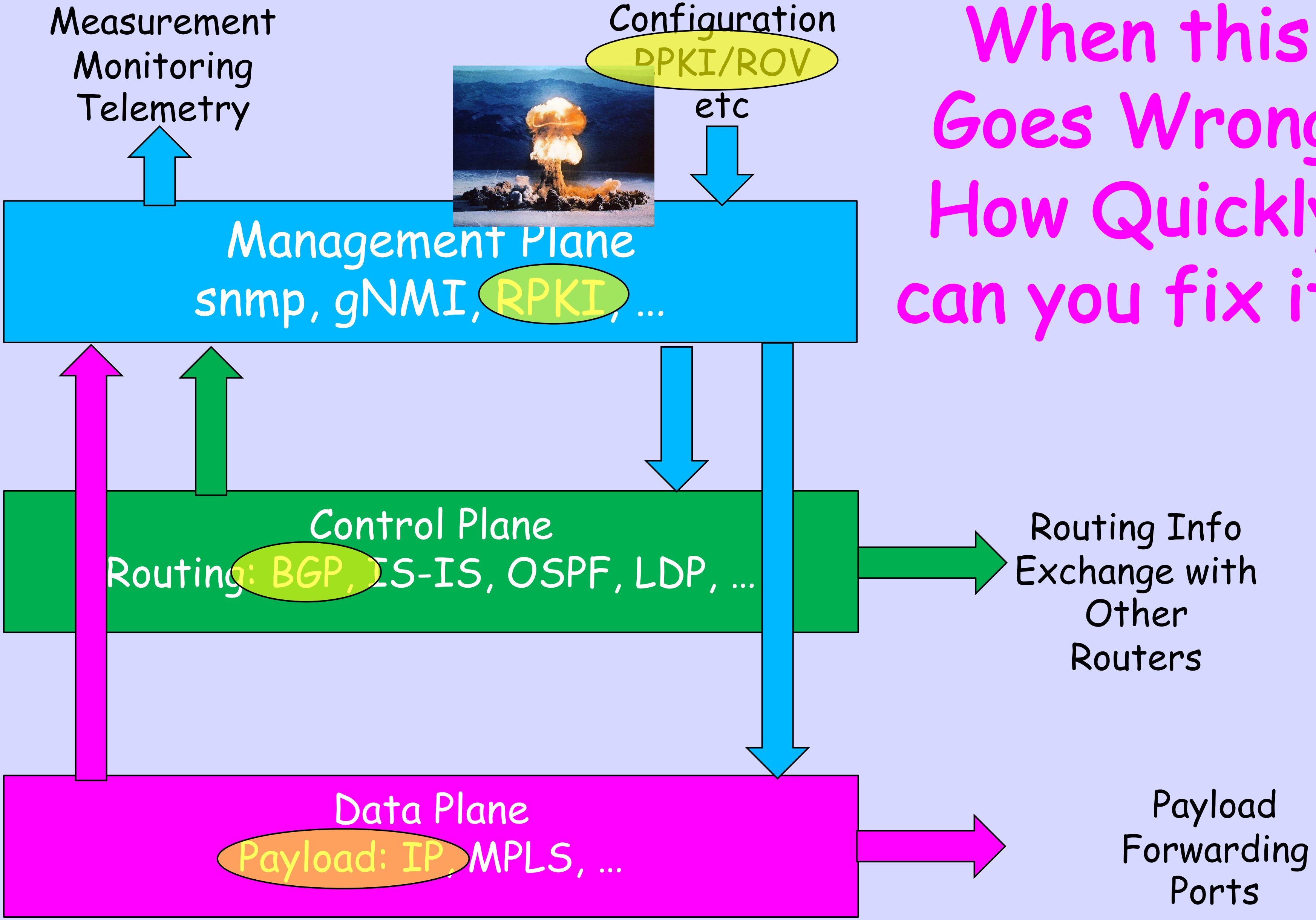
And ISP Delay Looks Bigger



	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

When this Goes Wrong How Quickly can you fix it?



Measurement
Monitoring
Telemetry

Configuration
RPKI/ROV
etc

Management Plane
snmp, gNMI, RPKI, …

Control Plane
Routing: BGP, IS-IS, OSPF, LDP, …

Routing Info
Exchange with
Other
Routers

Data Plane
Payload: IP, MPLS, …

Payload
Forwarding
Ports

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

Limitations of Study

- Relying Party software:
  - Fixed fetch rate so poor resolution
  - Only one RP software package used
- Did not measure RP to Router. But that is Notify driven so *should be* fast
- Did not measure delegated CAs
- RIR API/Screen-Scrape unreliable

What We Can Do?

- CAs/RIRs Publish Very Frequently
- RPs Poll Frequently, RRDP Please

As Protocol Designers

- BGP Transport is
  - Dangerously Shared Fate
  - Unordered, Reordering Guaranteed
- DNS does not handle Make Before Break

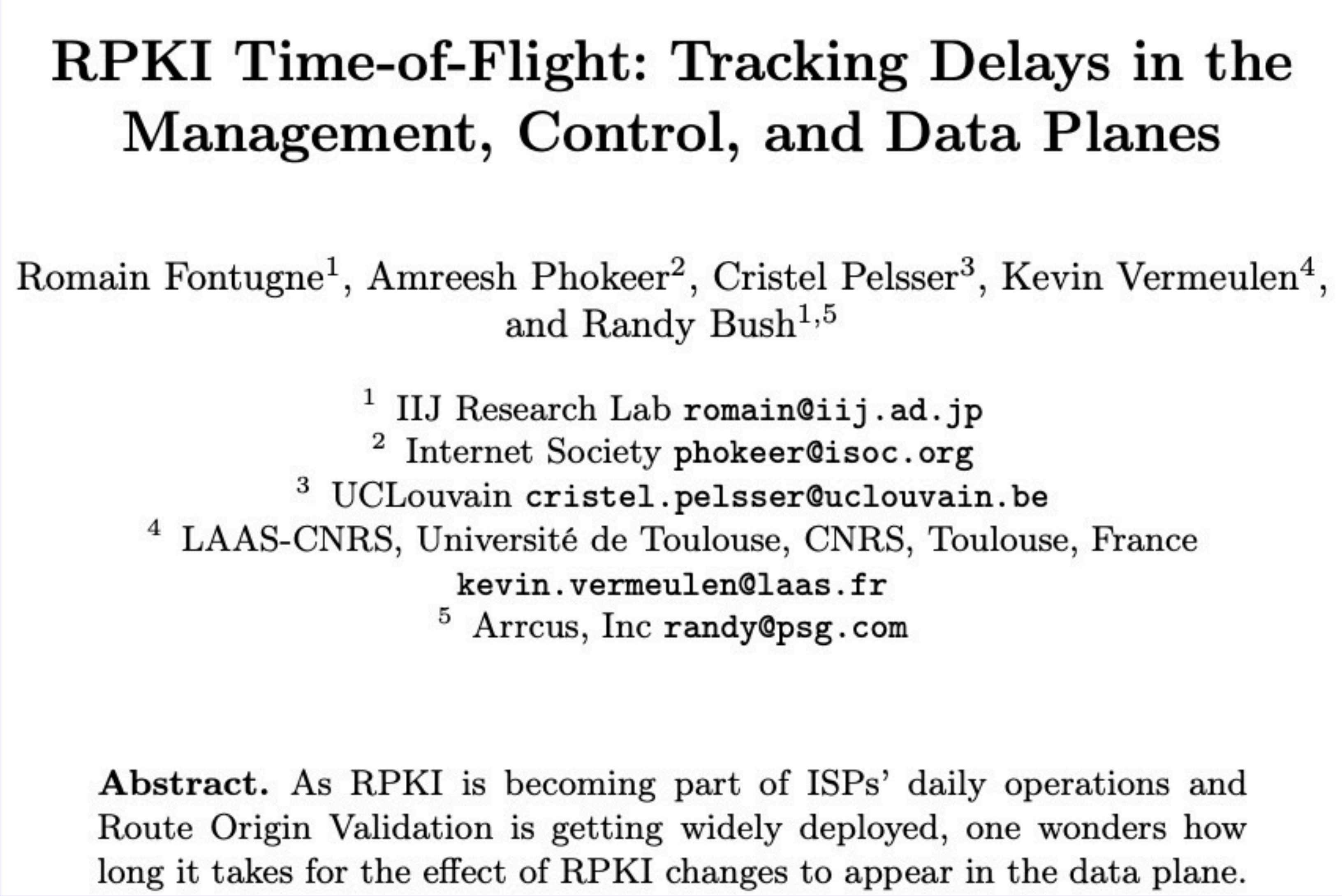
Is this Improving?

We Do Not Measure Our Networks Well ☹

Make Friends with a Measurement Researcher

☺

From the Paper in PAM 2023



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://archive.psg.com/pam2023-rov-ecosystem.pdf

Thanks To

Arrcus
Cisco
Equinix
Google
Juniper
NTT
Sprint

For Donated

- Rack Space
- Bandwidth
- Routers
- Switches
- Servers
- Etc. Etc.

Questions?

And Position Statements Pretending to be Questions ☺