

On Routing Security

Geoff Huston AM

Chief Scientist, APNIC Labs

Acknowledgement

- To George Michaelson for some of the material used here!

Geoff

Don't stop signing ROAs!

- This is **not** saying “RPKI is wrong and you shouldn't use it”
- Not at all!
- We have only a few tools to help us with keeping routing together, so we shouldn't let the perfect become the enemy of the good

Don't stop signing ROAs!

But

- If we can't be honest in appraising the effectiveness of these various approaches then we've walked away from evidence-based engineering and headed right into the fantasy marketing department!

Routing incidents comes in all shapes and sizes

- Some are malicious attacks intended to generate victims
- But most are incidents we inflict upon ourselves in various inept, random and accidental ways

What do *real* routing attacks look like?

Ethereum – 2018

- Effort that subverted the DNS, to take control over an Ethereum wallet
 - Subverted routing for Route 53 Authoritative DNS servers via more specific announcements for their own DNS server
 - This attack used a different origin AS (AS10297), but as it was a more specific prefix, they could've faked the original (AS16509) if we were doing origination checks at the time
 - Replied SERVFAIL for domains it wasn't attacking
 - Misdirected DNS for myetherwallet.com to intercepting website
 - If users clicked through a “this is a false certificate” warning, then their Ethereum wallet was drained
 - It was fast - the attack was all over in a couple of hours!
- The DNS attack could've been defeated by:
 - Users NOT clicking through a bad cert warning
 - Using DNSSEC signing for the myetherwallet.com domain
 - Using RPKI ROA with a maxlength parameter

What do *real* routing attacks look like?

Attacks tend to be multi-part these days

- Subvert the infrastructure enough to fool a DNS registrar
 - Take over the name registration and delegate the name
 - Re-sign the name with DNSSEC
 - Grab a cert from an CA
 - You're in!

Or

- Fool a CA's automated tests to get a fake certificate
 - By a targeted attack on the DNS resolution infrastructure
 - Then attack routing and use the fake cert to redirect users
 - You're in!

And then there are all the
rest...

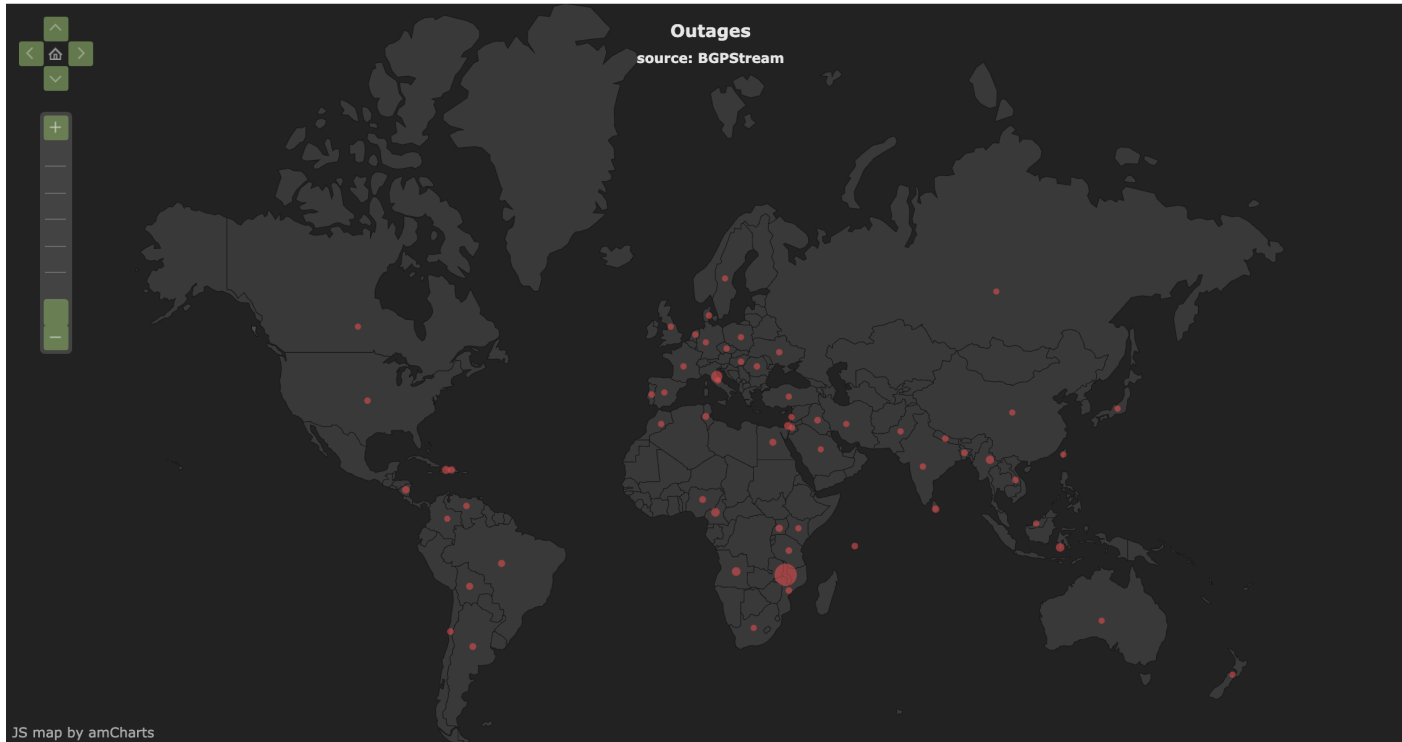
We get routing wrong a lot of the time



BGPmon is Now Part of **CrossworkCloud**

[Find Out More](#)

[BGPStream](#) [About](#) [Contact](#)



All Events for BGP Stream					
Event type	Country	ASN	Start time (UTC)	End time (UTC)	More info
Outage		Unknown (AS 58336)	2021-11-03 20:52:00	2021-11-03 20:55:00	More detail
Outage		Unknown (AS 58336)	2021-11-03 20:25:00	2021-11-03 20:32:00	More detail
Possible Hijack		Expected Origin AS: VITALITY_GB (AS 64456) Detected Origin AS: NTL_GB (AS 6089)	2021-11-03 20:04:35		More detail
Outage		ANONYMIZER, US (AS 53559)	2021-11-03 19:55:00		More detail
Outage		Unknown (AS 58336)	2021-11-03 19:02:00	2021-11-03 19:05:00	More detail
Outage		Unknown (AS 58336)	2021-11-03 18:01:00	2021-11-03 18:45:00	More detail
Outage		Unknown (AS 58336)	2021-11-03 17:51:00	2021-11-03 17:55:00	More detail
Outage		Turbo SP Internet Provider, BR (AS 52930)	2021-11-03 17:30:00		More detail
Outage		Unknown (AS 58336)	2021-11-03 17:22:00	2021-11-03 17:25:00	More detail
Outage		Unknown (AS 58336)	2021-11-03 17:07:00	2021-11-03 17:11:00	More detail
BGP Leak		Origin AS: VIETTEL-CAMBODIA-AS-AP-ISPXP IN CAMBODIA WITH THE BEST SERVICE IN THERE!, KH (AS 39620) Leader AS: VIETEL-AS-AP Viettel Group, VN (AS 7552)	2021-11-03 17:04:12		More detail
Outage		Unknown (AS 58336)	2021-11-03 16:53:00	2021-11-03 16:56:00	More detail
Possible Hijack		Expected Origin AS: --No Registry Entry-- (AS 64011) Detected Origin AS: Unknown (AS 64013)	2021-11-03 16:40:23		More detail
Outage		KNOU-AS Korea National Open University, KR (AS 10073)	2021-11-03 16:00:00		More detail
Outage		Unknown (AS 58336)	2021-11-03 15:42:00	2021-11-03 16:02:00	More detail
Outage		Unknown (AS 58336)	2021-11-03 15:27:00	2021-11-03 15:31:00	More detail
Outage		DNC-ASBLK-00306-00371, US (AS 336)	2021-11-03 15:22:00	2021-11-03 15:28:00	More detail
Outage		DNC-ASBLK-00306-00371, US (AS 336)	2021-11-03 15:21:00	2021-11-03 15:28:00	More detail
Outage		EDATEL TELECOMUNICACOES LTDA., BR (AS 262310)	2021-11-03 14:53:00		More detail
Outage		Unknown (AS 58336)	2021-11-03 14:46:00	2021-11-03 15:02:00	More detail
Outage		Unknown (AS 58336)	2021-11-03 14:10:00	2021-11-03 14:13:00	More detail
Outage		Unknown (AS 58336)	2021-11-03 13:56:00	2021-11-03 13:59:00	More detail
Outage		MICROLOGIC, US (AS 395209)	2021-11-03 13:55:00		More detail
Outage		Unknown (AS 58336)	2021-11-03 13:27:00	2021-11-03 13:30:00	More detail
Outage		L.M TIKO KAMIDE - SVA, BR (AS 28359)	2021-11-03 13:23:00		More detail
Outage		Unknown (AS 58336)	2021-11-03 12:30:00	2021-11-03 12:49:00	More detail
Outage	HT	N/A	2021-11-03 12:20:00		More detail
Outage		Unknown (AS 58336)	2021-11-03 12:16:00	2021-11-03 12:49:00	More detail
Outage		DNC-ASBLK-00306-00371, US (AS 346)	2021-11-03 12:10:00	2021-11-03 12:14:00	More detail
Outage		Unknown (AS 58336)	2021-11-03 11:44:00	2021-11-03 11:56:00	More detail
Outage		Unknown (AS 58336)	2021-11-03 11:21:00	2021-11-03 11:25:00	More detail
Outage		DNC-ASBLK-00306-00371, US (AS 346)	2021-11-03 11:00:00	2021-11-03 11:04:00	More detail
			2021-11-03	2021-11-03	More

And we would all like to handle this better

- So we take one or two proto-typical attacks and we design tools to prevent those attacks
 - It's simpler and focusses the effort to mitigate the issue
- And we hope that the ones we selected as attack examples were really good exemplars of what we are trying to prevent

The Archetypical BGP Incident

c|net

REVIEWS ▾

NEWS ▾

TECH ▾

MONEY ▾

WELLNESS ▾

HOME ▾

CARS ▾

DEALS ▾

February 2008

How Pakistan knocked YouTube offline (and how to make sure it never happens again)

YouTube becoming unreachable isn't the first time that Internet addresses were hijacked. But if it spurs interest in better security, it may be the last.

How?

- AS36561 (YouTube) was announcing 208.65.152.0/22
- AS17557 (Pakistan Telecom) announced 208.65.153.0/24

In BGP more specific prefixes “win” every time – so if a network heard the /24 then it believed it as a refinement to the encompassing /22

This was a failure in filtering.

But while (some) ISPs filtered their customers, the practise of applying filters to internal wholesale connections was less common. So the false route propagated from a regional transit provider and everyone else believed it.

But that was 14 years ago

- Are we getting better at filtering?
- Not really
- February 2021 AS 136168 (Campana MYTHIC) in Myanmar implements a government directive and propagates a more specific of Twitter's service address (104.244.42.0/24)
 - Twitter had lodged an RADB entry and if others were filtering on this RADB entry, then that would've stopped the false route propagating
 - But not enough folk perform transit filtering on RADB data
 - So the route propagated outward to AS132132, AS61292, AS4844, AS8106 and AS23673 and onward to ~40 other ASes who don't filter based on RADB entries

Maybe we need more than Route Registries...

- We've been using Route Registries as the foundation of route filtering since the NSF-funded Routing Arbiter project of the early 90's
- The problem with route registries is that they require intense feeding and watering, as they develop bitrot very quickly
- Surely we could use the Awesome Power of Digital Cryptography and automate the heck out of this and not just rely on hand-curated lists and fallible human operators?

Meanwhile, over the ditch in RIR Land

- We were looking at how to provide testable authenticity in supporting whois queries in the address registries
- The RIRs were aware that many ISPs used these registries as a source of authenticity to process requests to route BYO prefixes from customers, and ISPs were keen to push the authenticity problem off to literally anyone else!
- Maybe we could inject this testable authenticity directly into the routing system to literally make it impossible to lie!

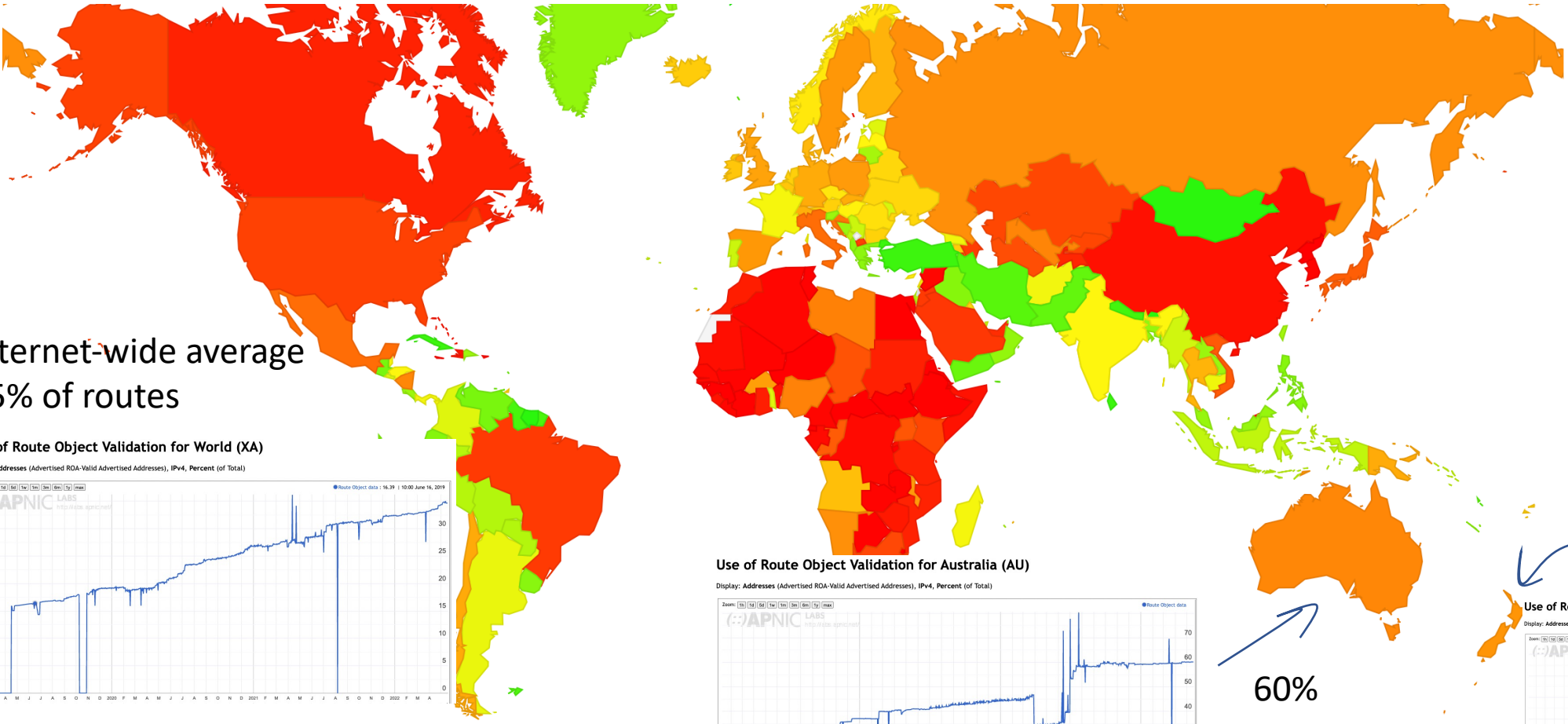
All Hail RPKI!

- Use the RIR registry as the source of authority
- Registry operator “certifies” a resource holder through a public key certificate
- Resource holders can digitally sign attestations with their resources
 - The signature also means that they are acknowledged resource holder and the resource is validly allocated or assigned
 - Validation of a signature means that the attestation is genuine, complete and current
- We can feed this into the routing system

RPKI, meet BGP!

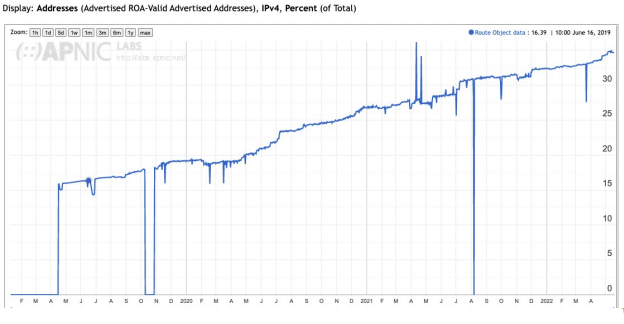
- Separate out *origination* and *propagation* and treat them separately
 - **Origination** is protected by using a signed authority issued by the prefix holder to authorise an AS to originate a route or set of routes (ROA)

ROA Production

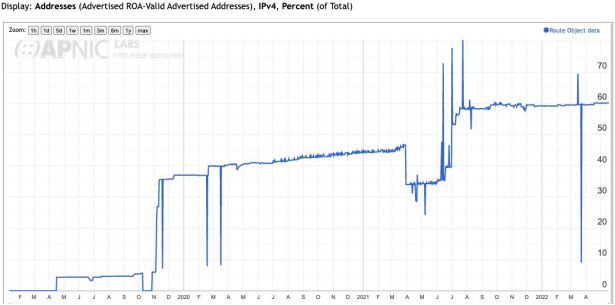


Internet-wide average
35% of routes

Use of Route Object Validation for World (XA)

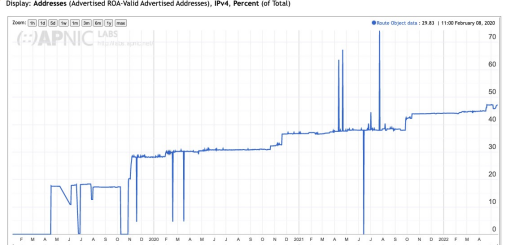


Use of Route Object Validation for Australia (AU)



60%

Use of Route Object Validation for New Zealand (NZ)



45%

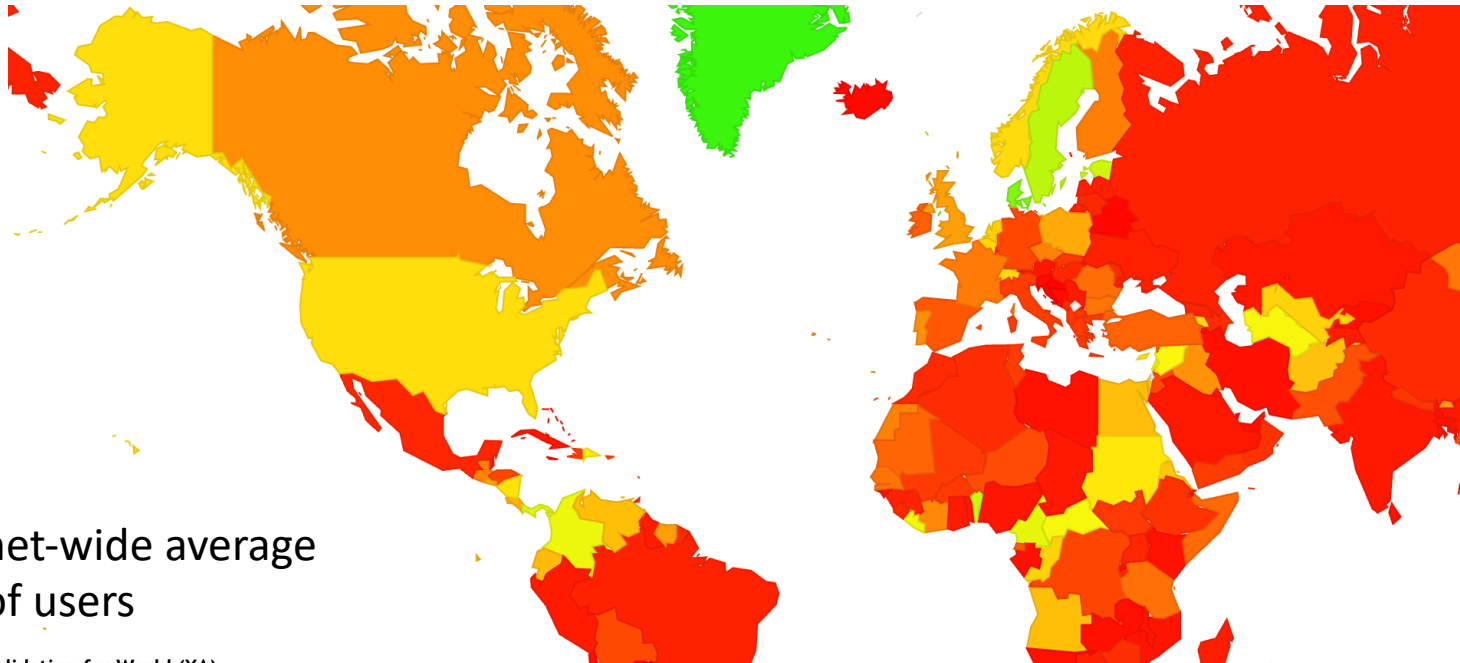
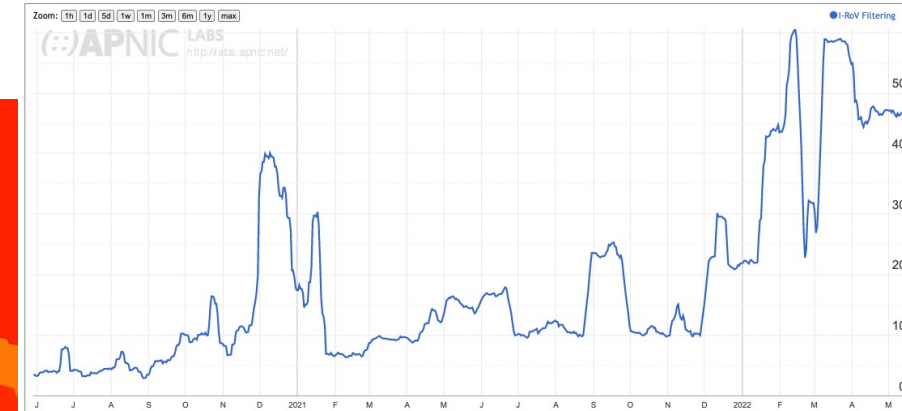
<https://stats.labs.apnic.net/ROAS>

RPKI, meet BGP!

- Separate out *origination* and *propagation* and treat them separately
 - **Origination** is protected by using a signed authority issued by the prefix holder to authorise an AS to originate a route or set of routes (ROA)
 - Then you assemble these digital authorities and generate a filter list in the router and drop all routing announcements that are invalid

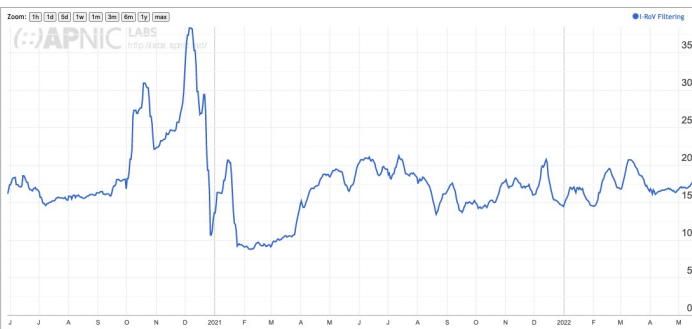
Drop RoV-Invalid routes

Use of RPKI Validation for New Zealand (NZ)

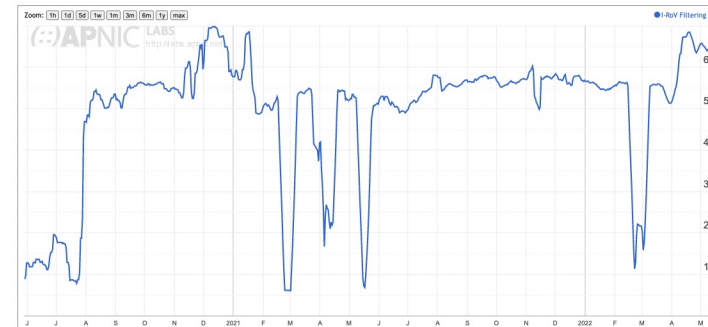


Internet-wide average
16% of users

Use of RPKI Validation for World (XA)



Use of RPKI Validation for Australia (AU)



45% of users
65% of users
100

RPKI, meet BGP!

- Separate out *origination* and *propagation* and treat them separately
 - **Origination** is protected by using a signed authority issued by the prefix holder to authorise an AS to originate a route or set of routes (ROA)
 - Then you assemble these authorities and generate a filter list in the router and drop announcements that are invalid
 - **Propagation** is a problem
 - Wholistic approaches that attempt to link the AS path to the propagation of an update (BGPSEC) resist piecemeal deployment and are crypto-intensive – BGPSEC is largely DOA
 - Piecemeal approaches offer more limited protections that limit the plausibility of some forms of lies

RPKI, meet BGP!

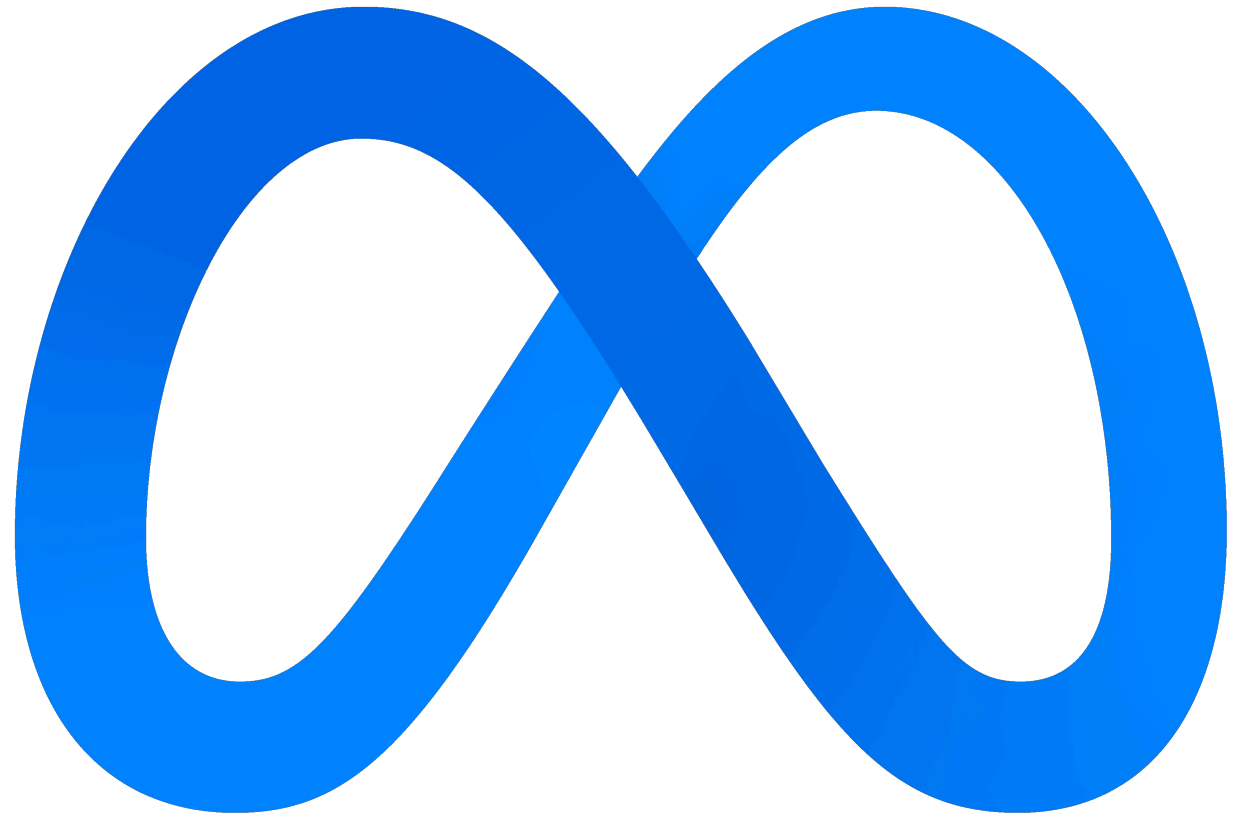
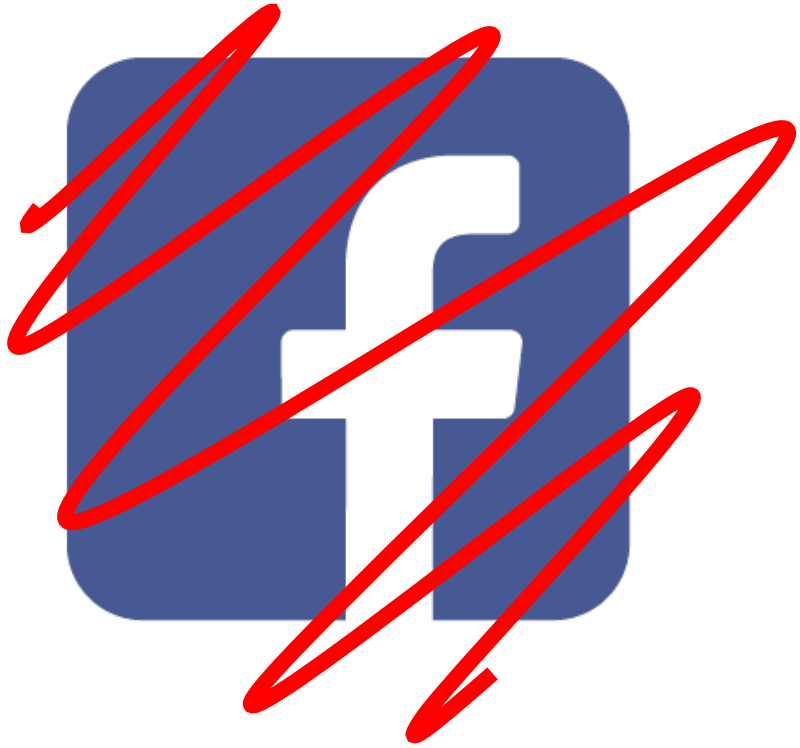
- It's not an entirely comfortable match between RPKI and BGP
- The hop-by-hop aspects of BGP (withdrawals, communities) are not possible to validate against an origination "root cause"
- Routing is "backwards"
 - BGP does NOT select the forwarding path
 - It creates a partial topology by passing reachability in the reverse direction
 - And that's all
 - An AS Path describes the route propagation path, not the packet's forwarding path
- What matters is "forwards"
 - Our concern is with the forwarding path
 - And that's what we can't check from the routing system

So securing routing is hard

But is it enough?

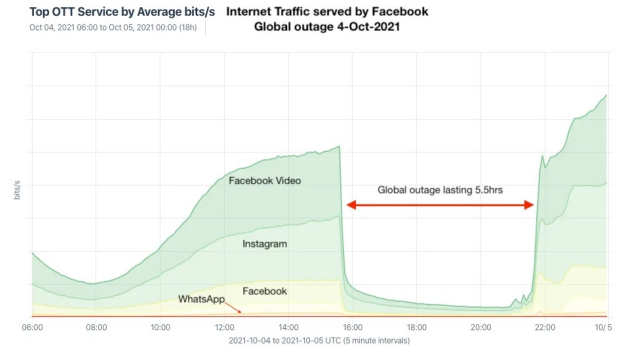
- What do we see in terms of “incidents” in today's network?
- Would RPKI in BGP defend the network from such incidents in any case?
- Let's look at some more examples of incidents and outages that induced routing anomalies





~~Facebook~~ Meta, October 2021

- Lost route to nameservers
 - Used a single route to all nameservers
 - Used short cache lifetime DNS records
 - Lost sight of all 4 anycast nameservers for @facebook.com
- Lost access to secure entry tokens in @facebook.com
- Even if they'd had DNS, NLRI routes to offline server surface would have looked pretty bad
 - Web service sending what?
- 6-8 hour outage
 - (5nines is 5min/year unplanned outage, so that's 72 years of 'credit' for 5nines!)
- Huge amounts of Africa functionally offline for business
 - WhatsApp for money exchange
- 3 billion Facebook users were totally disconnected from the platform over this time



Own Goal Syndrome




Yes, there was a BGP incident at the heart of this

- It was a withdrawal that isolated the authoritative nameservers for facebook.com
- But it was not an attack
- It was an internal operational error
- And RPKI/BGPSEC cannot “protect” inadvertent route withdrawals in any case
- And the outage was multiplied by the withdrawal of the DNS records because of cache expiry
- And made worse because the outage also locked them out of their facilities (!)

More recent Own Goals



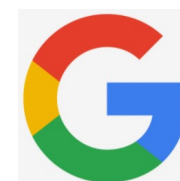
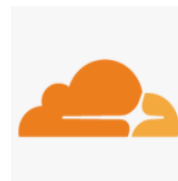
- **fastly** June 2021
 - A certain customer configuration change that was flagged as valid triggered a complete platform crash in their Varnish platform
 - Varnish is NOT a Fastly-developed platform – it is open source developed by a Norwegian newspaper site
-  **Akamai** July 2021
 - A config change had a format error that disabled the front end load balancer, that disabled their DNS steering and took out the platform
 - Obviously, not a routing problem

Own Goals are insanely common!

- But what about collateral damage?

Collateral Damage

- November 2018, MainOne in Nigeria had a configuration error that leaked ~200 Google Cloud routes to various transits, including China Telecom, who propagated the routes onward
- Two hours later Main One Leaked Cloudflare routes along the same transit paths
- Would RPKI have helped here?
 - Assuming this was a path leak, then no, not really
 - It's a routing policy violation, not a protocol / announcement correctness issue



Collateral Damage

- June 2019 AS33154 (DQE Communications, US) had a Noction BGP Optimizer that announced a set of more specifics to its customer AS396531 (Allegheny Technologies) who readvertised these routes to AS7012 (Verizon) a Large Tier 1 transit network who was not performing route filtering
- A large set of routes were redirected along this mule track detour, including AWS and Cloudflare, causing major disruption
- RPKI? Maybe – depends on the use of ROAs with maxlength to reject more specifics

The Verizon logo, consisting of the word "verizon" in a bold, black, sans-serif font with a red checkmark above the "i", all contained within a white rectangular box with a thin grey border and a subtle drop shadow.

verizon[✓]

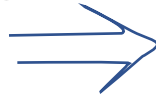
The AWS logo, featuring the letters "AWS" in a bold, white, sans-serif font centered within a white cloud shape, which is itself centered within an orange square, all contained within a white rectangular box with a thin grey border and a subtle drop shadow.

AWS



Other Recent Collateral Damage incidents

- IBM cloud outage, June 2020
 - “external provider leak”
- Unnamed external provider, 2+ hours, multiple regions.



- RPKI? Possibly, possibly not!

More

- TWC, Rogers, Charter, July 2020



- Small ISP deploying “BGP optimisers” leaked routes
 - Propagated by Telia

Yes, More

- Vodafone India prefix leak, April 2021
 - 30,000 prefixes mistakenly leaked
 - Google, Akamai, Edgecast, Deutsche Telekom, TIM, Claro, Orange, Telefonica), Vodafone itself (worldwide) amongst others



Commonalities

- Its an indirect sideswipe
 - But there is still a service loss, loss of business, customer/SLA effects
- Yes, some of these could have been avoided by a ROA with careful use of maxlength
 - But
 - ROAs are universal, not context specific
 - They don't come with an "apply here, but not there, sticker"
 - If ROAs allow you to accept more specifics, then they won't stop you propagating them onward
- But other incidents are policy routing issues relating to leakage, not synthetic routes

Does Network Automation help here?

- Yes - and No
- A config change can flip the state of all components of the network all at once – amplifying the potential for a problem to be network wide though just one command transaction
- Less margin for error and greater potential for damage.
- And automated scripts within these system can generate completely unanticipated outcomes!
- Some network managers see the purchase of an automated network management system as a compatible substitute for a skilled workforce
 - a stunning triumph of unwarranted optimism over reality!

Does RPKI help here?

Well, yes, a bit, but its not the full picture:

- And adding more moving parts to a complex system does not make it more robust – it often achieves the exact opposite
- RPKI uses a single rule set that is applied everywhere – it does not provide context-specific conditional application
- Many route leaks are a policy violation, not a protocol violation
 - And policies are often contextual, not universal
- Some of the routing issues are the result of loss of a synchronised forwarding state, and BGP (and RPKI) don't and can't enforce synchronicity in state across BGP speakers
 - We've seen “ghost routes” in BGP that have been persistent for years!

So... what's the answer?

- We continue to push larger route sets and larger policy agendas onto the routing system
- And because clue is finite, we are automating more and more of network management to make up for a serious skill gap
- Which creates brittleness in the routing that is prone to fail in unsafe states that can't be readily recovered
- How can we make routing more robust?

Routing is Difficult

"The most complicated computation ever attempted by mankind is the global distributed routing algorithm that runs the Internet.

In fact, if anybody thought about it very hard, before we started, they would've been too scared to try.

Ah, because it runs in near real-time, it's an online algorithm, it runs on a multimillion node multicomputer, of an arbitrary topology, built by lots of people who have never met each other. Right?

And, it's a very very complex computation because it's piecewise constructive, there is a lot of local consistency constraints, there is a bunch of global correctness criteria that are occasionally satisfied, and yet the thing mostly works.

Which is astounding, when you actually look at what's going on."

Routing Security is not a solved problem

- I'm not sure we really know what we really mean when we talk of routing security
- And I'm not sure that operationally focussed piecemeal incrementalism is really helping here with the bigger picture of tackling "routing robustness" and stopping these various routing mishaps
- This is remains a problem space that would benefit from further research and experimentation
- And research funding of course! 😊

Thanks!