

Routing Security is  
more than RPKI

Geoff Huston AM  
Chief Scientist, APNIC

# Don't stop signing ROAs!

- This is **not** saying RPKI is wrong and you shouldn't use it
- 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
  - Unless of course the entire Route Refresh story scares you, in which case turning on RoV might not be your best idea you've had today!*

# Don't stop signing ROAs!

- This is **not** saying RPKI is wrong and you shouldn't use it
- 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
  - Unless of course the entire Route Refresh story scares you, in which case turning on RoV might not be your best idea you've had today!*

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 and accidental ways

# What do real routing attacks look like?

## Ethereum – 2018

- Deliberate attempt to subvert DNS, to take over Ethereum wallets
  - 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 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 false SSL warning, then their Ethereum wallet was stolen
- This 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?

Lessons:

## **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!
- Fool the CA's 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!

## **Attacks are executed very quickly**

- From the original subversion to trigger the attack to completion is often less than a couple of hours
- Anything after that is a bonus

## **Attackers don't bother to clean up afterwards**

- Attackers don't care about certificate transparency, system logs or any other related forensic information
- They rely on existing network capabilities to hide their identities

And then there are all the  
rest...

- While these attacks get headlines, there are more mundane self-harm incidents in routing that happen all the time

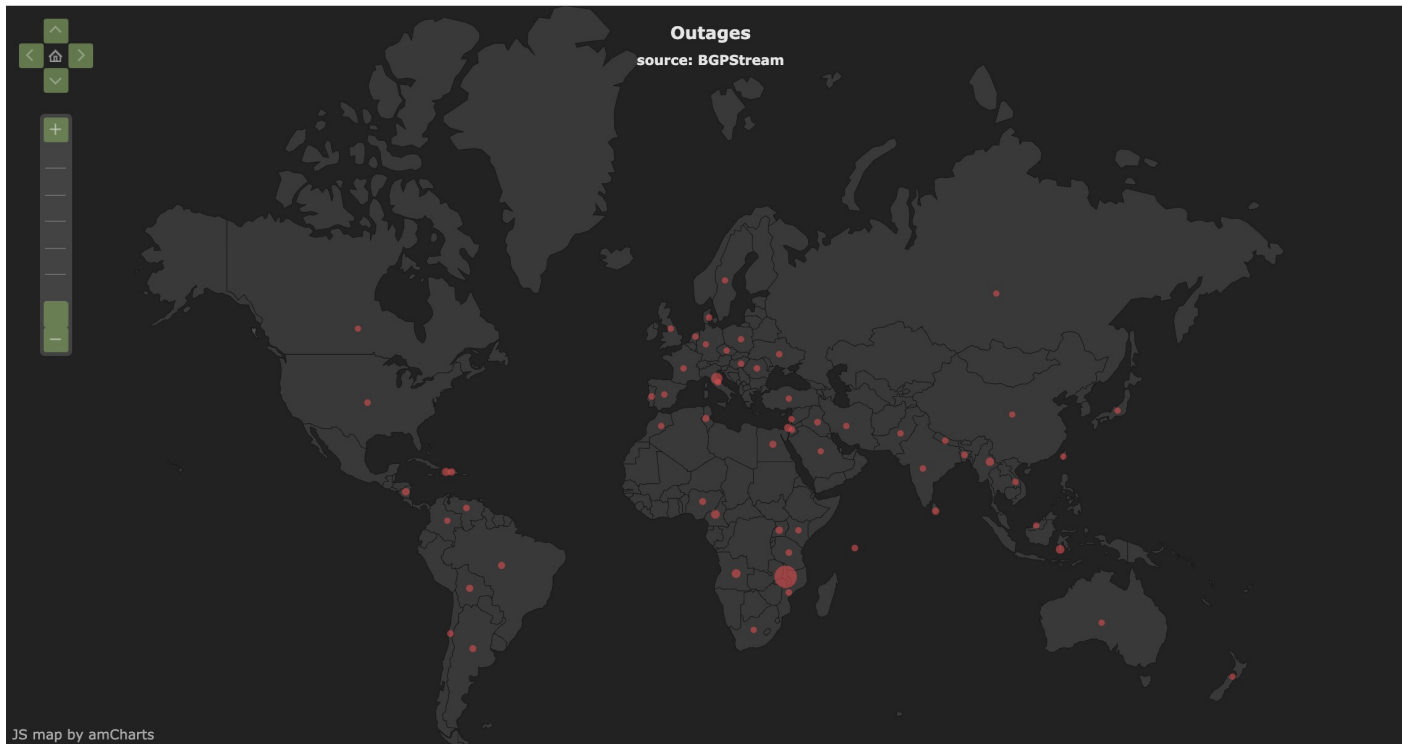
# We get routing wrong a lot of the time



BGPMon is Now Part of  
**CrossworkCloud**

[Find Out More](#)

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

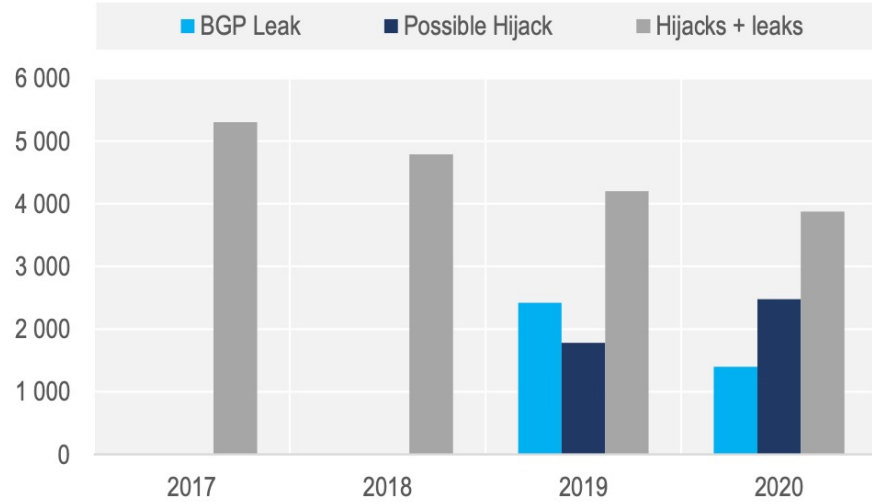


All Events for BGP Stream					
Event type	Country	ASN	Start time (UTC)	End time (UTC)	More info
Outage		Unknown (AS 58336)	2021-11-03 20:52:00	2021-11-03 20:55:00	<a href="#">More detail</a>
Outage		Unknown (AS 58336)	2021-11-03 20:25:00	2021-11-03 20:32:00	<a href="#">More detail</a>
Possible Hijack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

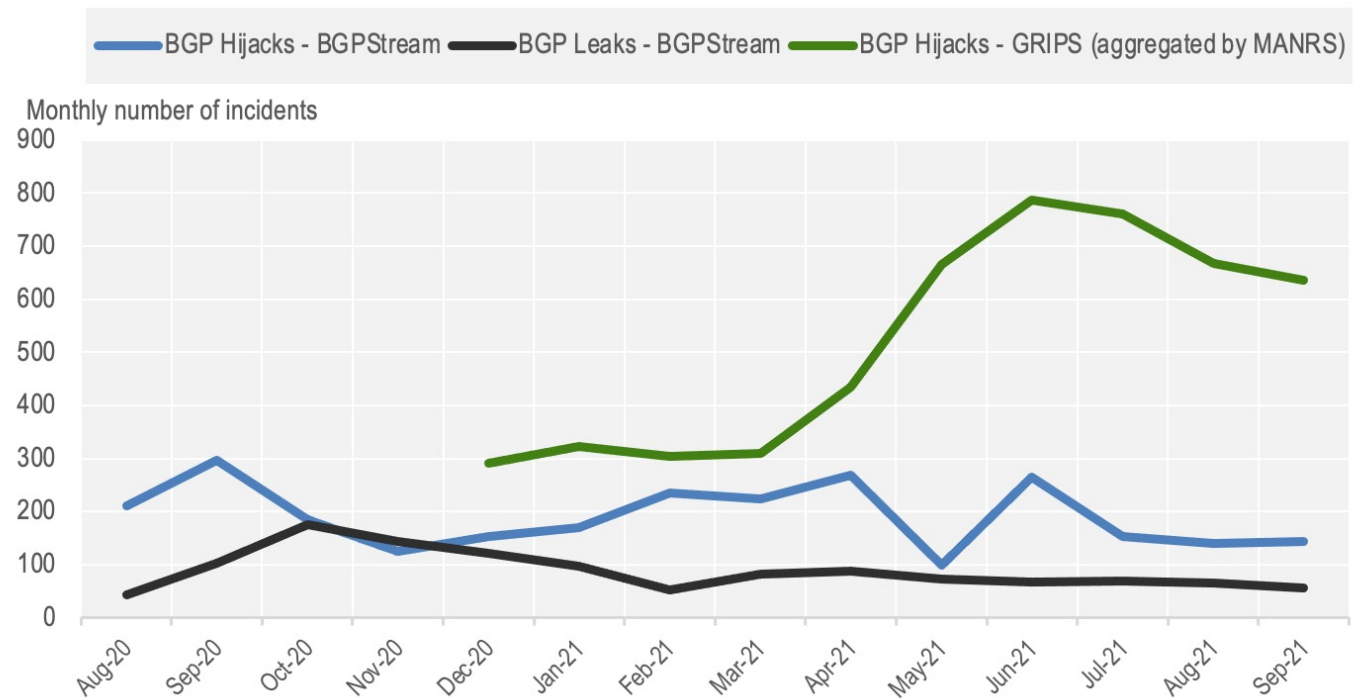
BGPMon report



# We get routing wrong a lot of the time



Annual summary of BGPStream reports



# And we would all like to handle this better

- So we take one or two proto-typical attacks
  - It's simpler and focusses the effort to mitigate the issue
- We hope that they are good examples of what we are trying to prevent
- We design tools to prevent those attacks

# 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 and everyone else believed it.

# But that was 13 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 filtering on this RADB entry would've stopped the false route propagating
  - Yet the route still 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 fence 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 information into the routing system to audit announcements for veracity



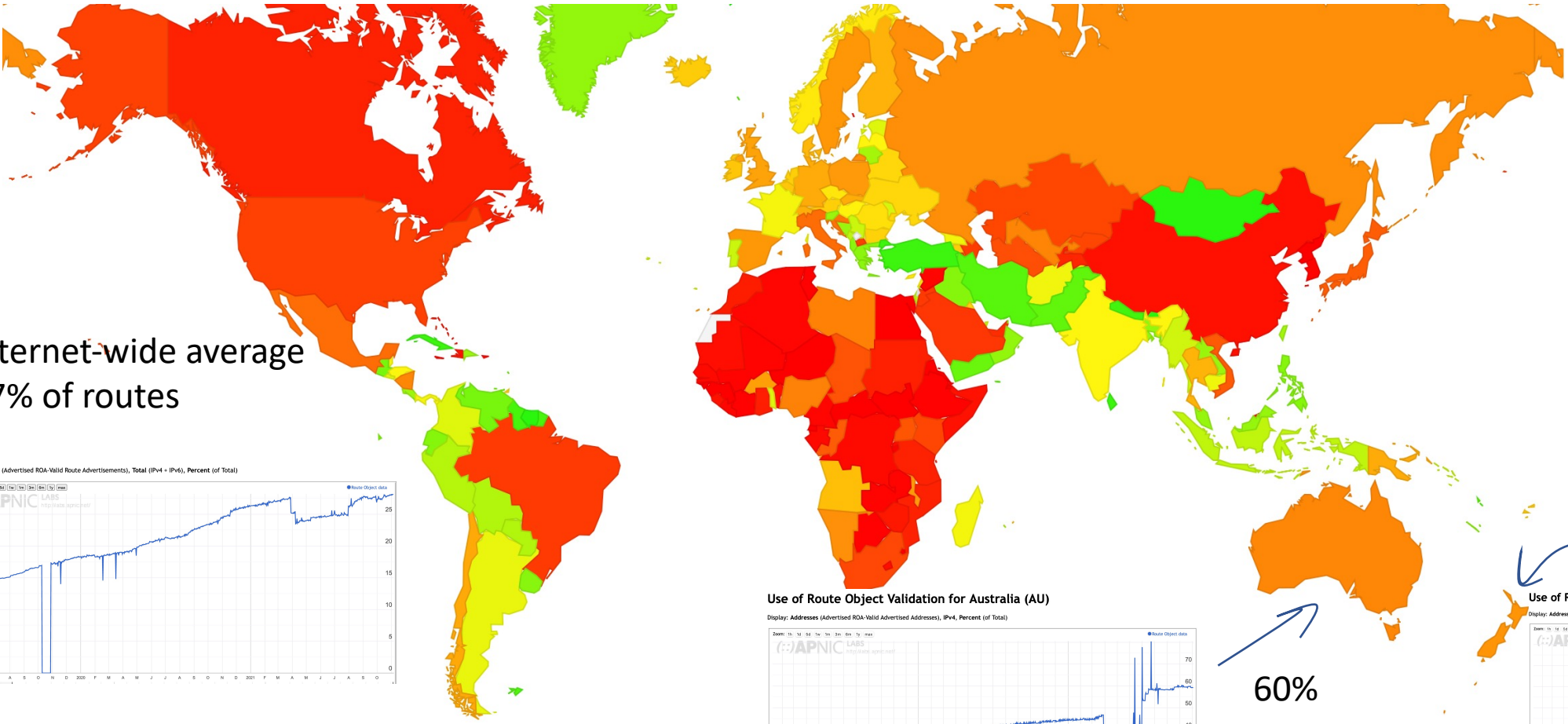
# RPKI, meet BGP!

- Separate out *origination* and *propagation* and treat them separately

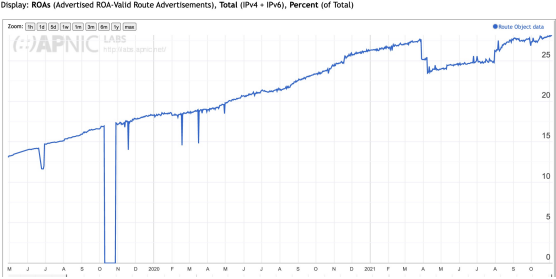
# RPKI, meet BGP!

- 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)

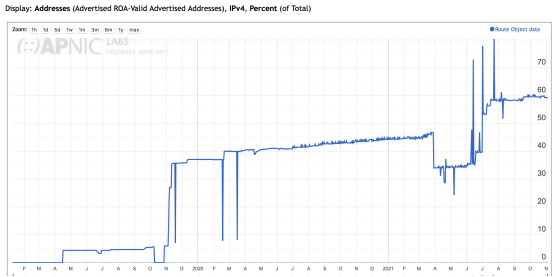
# ROA Production



Internet-wide average  
27% of routes

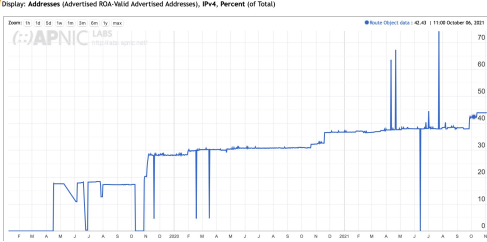


Use of Route Object Validation for Australia (AU)



60%

Use of Route Object Validation for New Zealand (NZ)

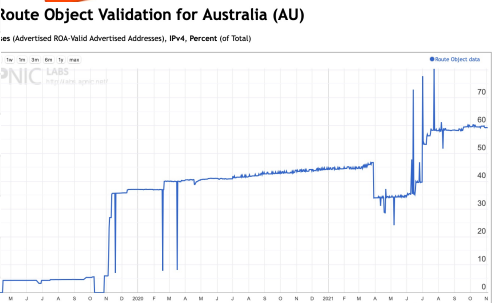
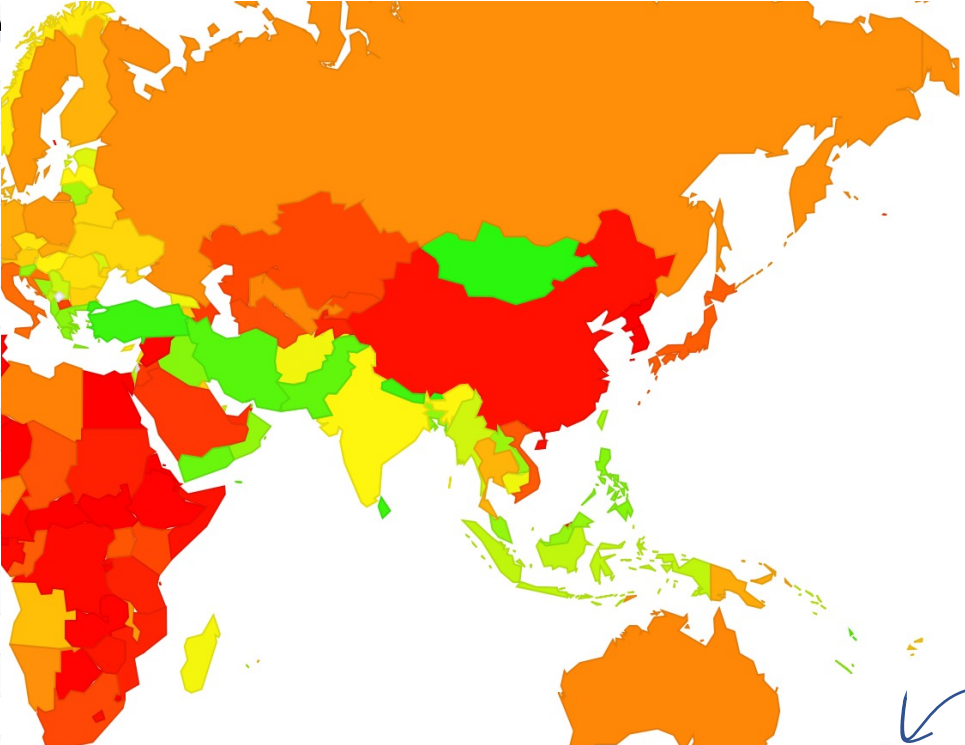


40%

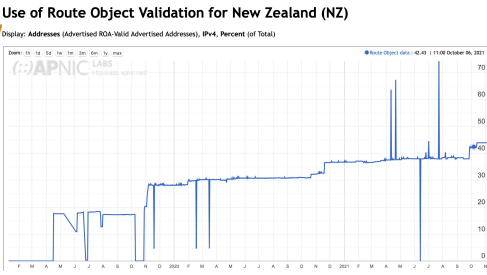
<https://stats.labs.apnic.net/ROAS>

# ROA Production

Code	Region	V4 Valid	Pc
NU	Niue, Polynesia, Oceania	512	100.0%
TO	Tonga, Polynesia, Oceania	9,472	97.4%
VU	Vanuatu, Melanesia, Oceania	14,080	90.2%
MH	Marshall Islands, Micronesia, Oceania	4,608	90.0%
NC	New Caledonia, Melanesia, Oceania	125,184	84.9%
FJ	Fiji, Melanesia, Oceania	117,504	83.0%
WS	Samoa, Polynesia, Oceania	16,384	79.0%
SB	Solomon Islands, Melanesia, Oceania	8,960	70.0%
AU	Australia, Australia and New Zealand, Oceania	26,582,113	58.0%
NR	Nauru, Micronesia, Oceania	5,632	57.9%
NZ	New Zealand, Australia and New Zealand, Oceania	2,494,976	44.0%
PG	Papua New Guinea, Melanesia, Oceania	16,384	23.4%
GU	Guam, Micronesia, Oceania	30,208	13.6%
PF	French Polynesia, Polynesia, Oceania	8,448	12.5%
KI	Kiribati, Micronesia, Oceania	512	11.1%
MP	Northern Mariana Islands, Micronesia, Oceania	1,024	3.7%
AS	American Samoa, Polynesia, Oceania	512	3.0%
CK	Cook Islands, Polynesia, Oceania	0	0.0%
FM	Micronesia (Federated States of), Micronesia, Oceania	0	0.0%
NF	Norfolk Island, Australia and New Zealand, Oceania	0	0.0%
PW	Palau, Micronesia, Oceania	0	0.0%
TK	Tokelau, Polynesia, Oceania	0	0.0%
TV	Tuvalu, Polynesia, Oceania	0	0.0%
WF	Wallis and Futuna Islands, Polynesia, Oceania	0	0.0%



60%

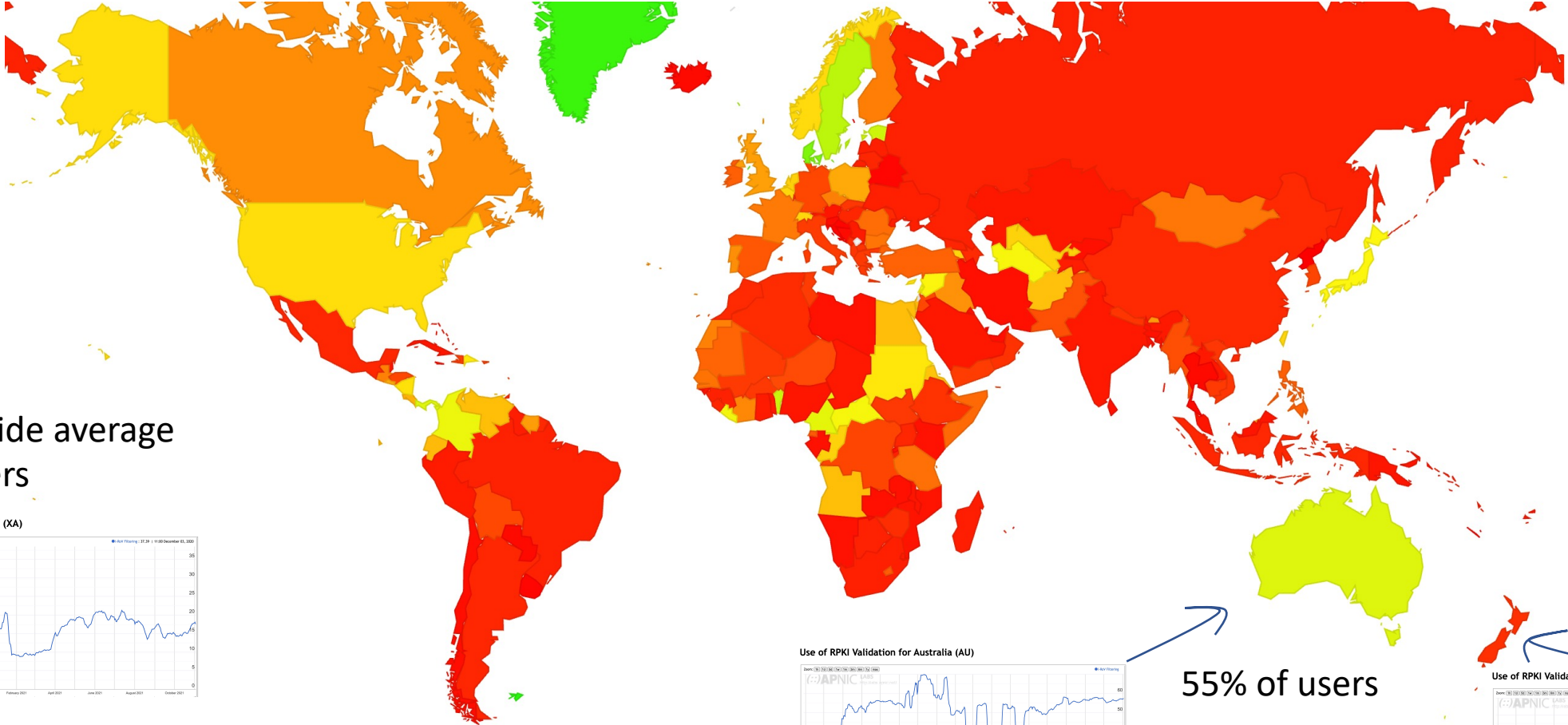


40%

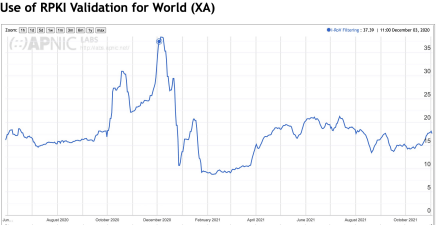
# RPKI, meet BGP!

- 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

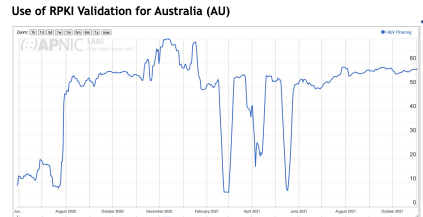
# Drop RoV-Invalid routes



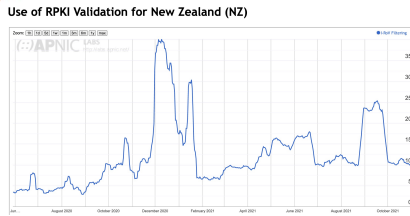
Internet-wide average  
15% of users



55% of users

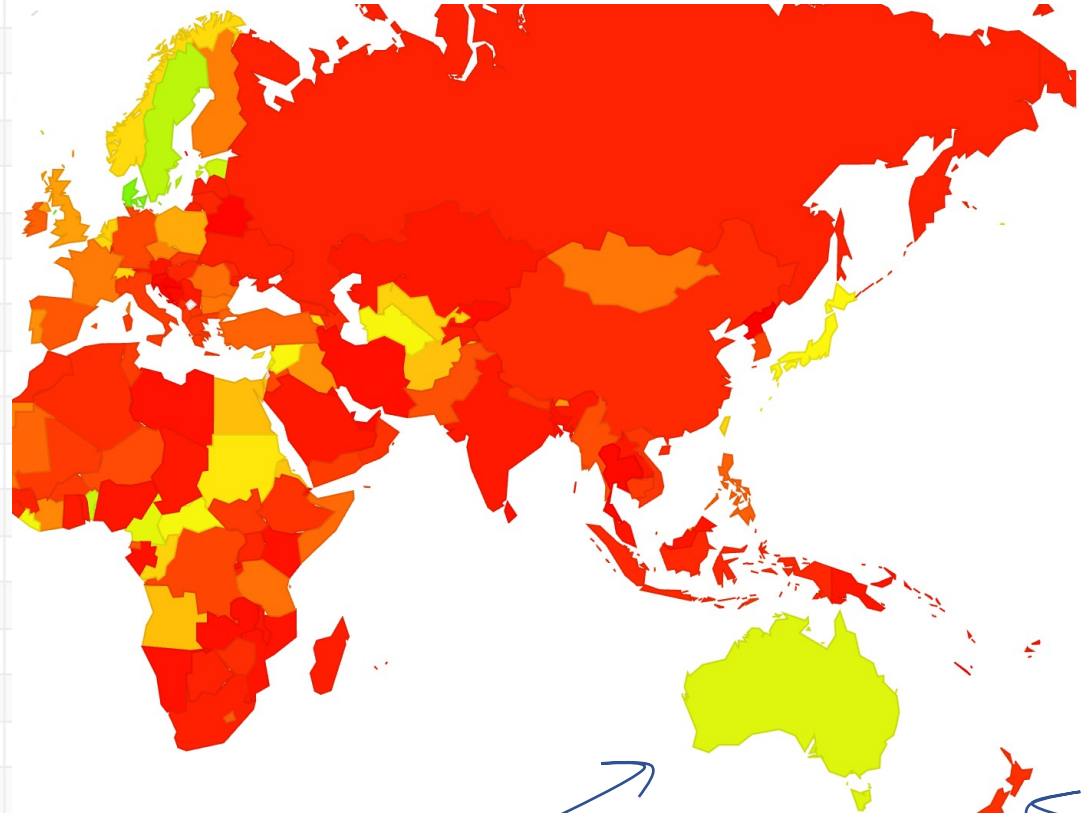


10% of users

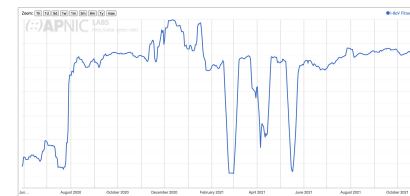


# Drop RoV-Invalid routes

CC	Country	RPKI Validates
MH	Marshall Islands, Micronesia, Oceania	82.43%
TV	Tuvalu, Polynesia, Oceania	62.99%
AU	Australia, Australia and New Zealand, Oceania	56.28%
WF	Wallis and Futuna Islands, Polynesia, Oceania	48.61%
TO	Tonga, Polynesia, Oceania	46.09%
WS	Samoa, Polynesia, Oceania	35.95%
KI	Kiribati, Micronesia, Oceania	21.43%
CK	Cook Islands, Polynesia, Oceania	18.60%
PF	French Polynesia, Polynesia, Oceania	17.68%
AS	American Samoa, Polynesia, Oceania	17.25%
NF	Norfolk Island, Australia and New Zealand, Oceania	13.04%
NZ	New Zealand, Australia and New Zealand, Oceania	11.52%
PG	Papua New Guinea, Melanesia, Oceania	9.52%
NR	Nauru, Micronesia, Oceania	8.74%
MP	Northern Mariana Islands, Micronesia, Oceania	7.27%
FM	Micronesia (Federated States of), Micronesia, Oceania	7.09%
SB	Solomon Islands, Melanesia, Oceania	6.90%
FJ	Fiji, Melanesia, Oceania	5.26%
VU	Vanuatu, Melanesia, Oceania	3.82%
GU	Guam, Micronesia, Oceania	3.09%
PW	Palau, Micronesia, Oceania	2.63%
NC	New Caledonia, Melanesia, Oceania	1.32%

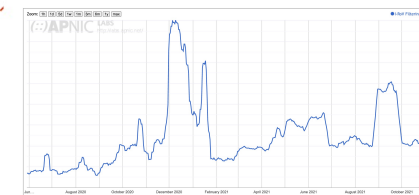


Use of RPKI Validation for Australia (AU)



55% of users

Use of RPKI Validation for New Zealand (NZ)



10% of users

# 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**



# 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

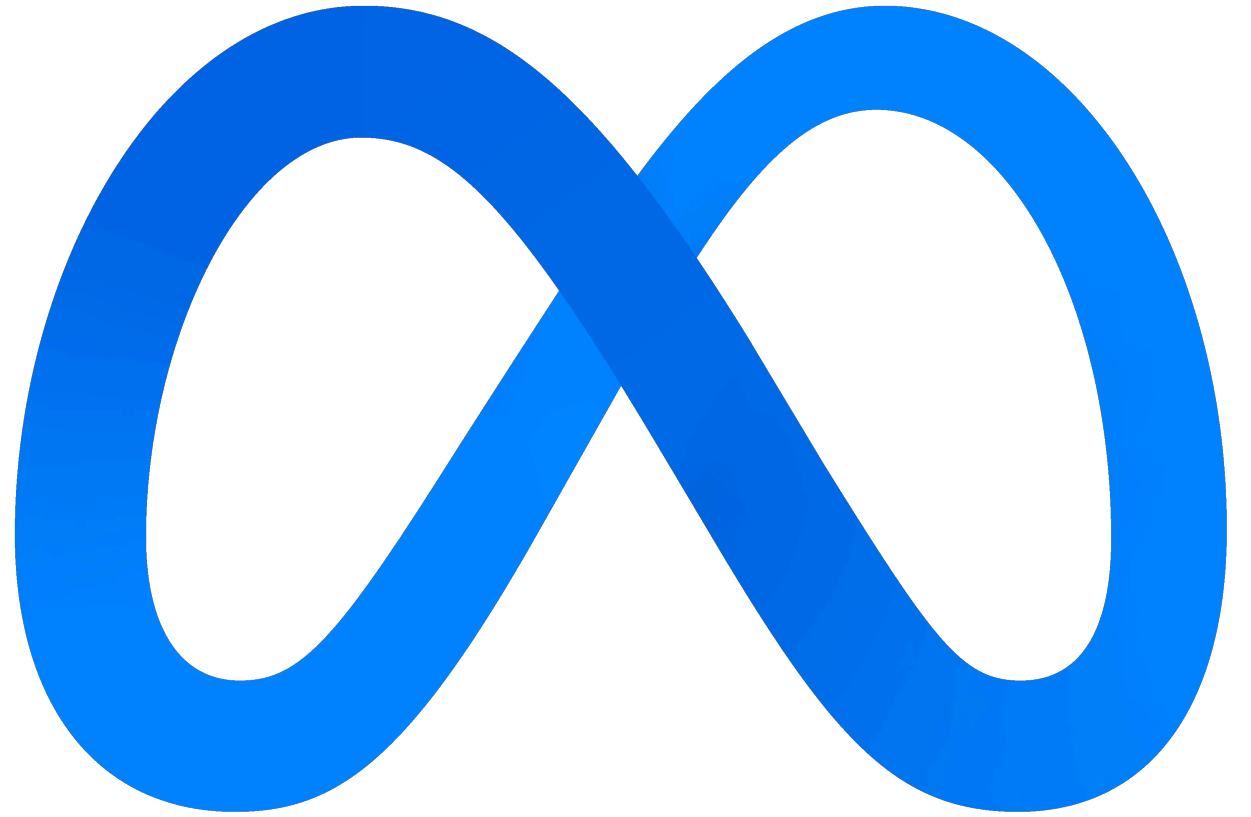
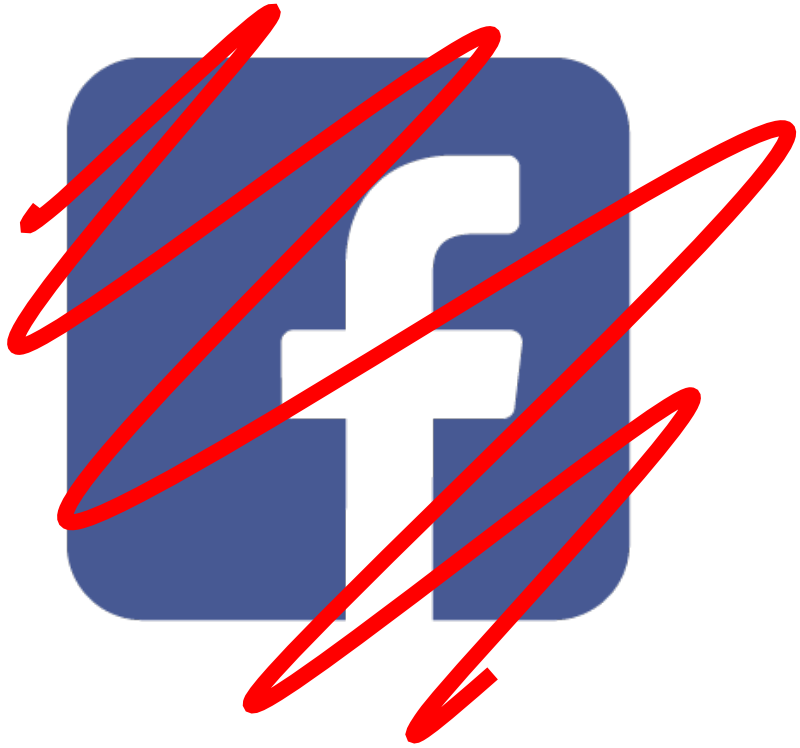
- The implicit withdrawal of ROA-invalid routes can cause spurious route refresh operations against your BGP peers
- The hop-by-hop aspects of BGP (withdrawals, communities) are not possible to validate against an origination “root cause”
- Routing is “backwards”
  - BGP does NOT select the forwarding 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 its still only part of the picture

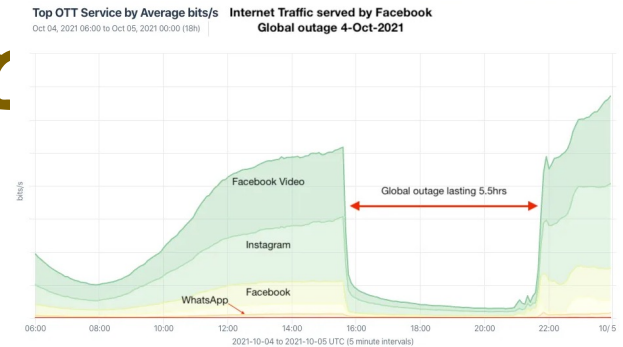
- What do we see in terms of “incidents” in todays network?
- Would RPKI in BGP defend the network from such incidents in any case?
- Let’s look at some more examples of incidents and outages that induced routing anomalies





# ~~Facebook~~ Meta, October

- 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 totally disconnected from the platform



# Own Goal Syndrome




Yes, there was a BGP incident

- 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

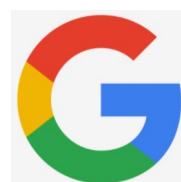
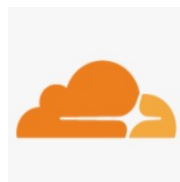


# Own Goals are insanely common!

- 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



# 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



verizon<sup>✓</sup>



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!

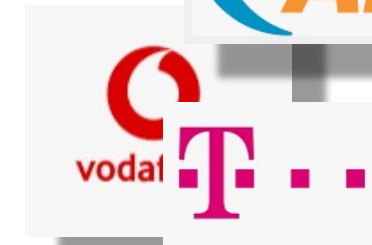
# More

- 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 many 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 sometimes, 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
  - So RPKI won't catch them
- 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! 😊

# But there really are some things you should do in routing

- SECURE YOUR ROUTERS!
- Avoid multi-hop EBGP wherever possible
  - And if you must multi-hop, use TCP-AO or MD5 to secure the channel
- Use an IRR
  - Yeah, it may seem like a bit of a waste of time, but honestly the rigor of enumerating your routes and keeping these records up to date is worth it even if you are the only user of the IRR entry!
- Apply Source filters to prevent source address spoofing
- Look at your network from afar – all of the time!
- Look at your routes from afar – all of the time!

# More BGP Safety First

- Rehearse every routing config change
- Always have a backout plan prepared
- Talk with your BGP peers, upstreams and customers
  - So at the very least you know how to reach them when you are desperate!
- Don't try and cover up any routing mistakes – be honest and open so that we can all help to clean it up as quickly as possible
  - And yes, we've all been responsible for our share of routing lapses
- Certify your resources
- Sign ROAs
- Filter inbound and outbound routes when you (and your router) feel ready!
- Follow the technology yourself
  - Reliance on vendors to do all the heavy lifting for you is generally a mistake

# More safety first

Its not directly related to routing, but it really helps your overall security stance

- Use a DNSSEC-validating recursive resolver to server your customers
  - YES, all of the time!
  - No exceptions
- DNSSEC-sign your domain names
  - Because the worst attacks corrupt the DNS as well as routing

# It won't catch everything

- But hopefully you will feel more confident that you can manage your network to provide a stable service
- And you will be more confident that you can recover quickly when things don't quite go as planned!



Thanks!