1.1.1.0/24
A report from the (anycast) trenches

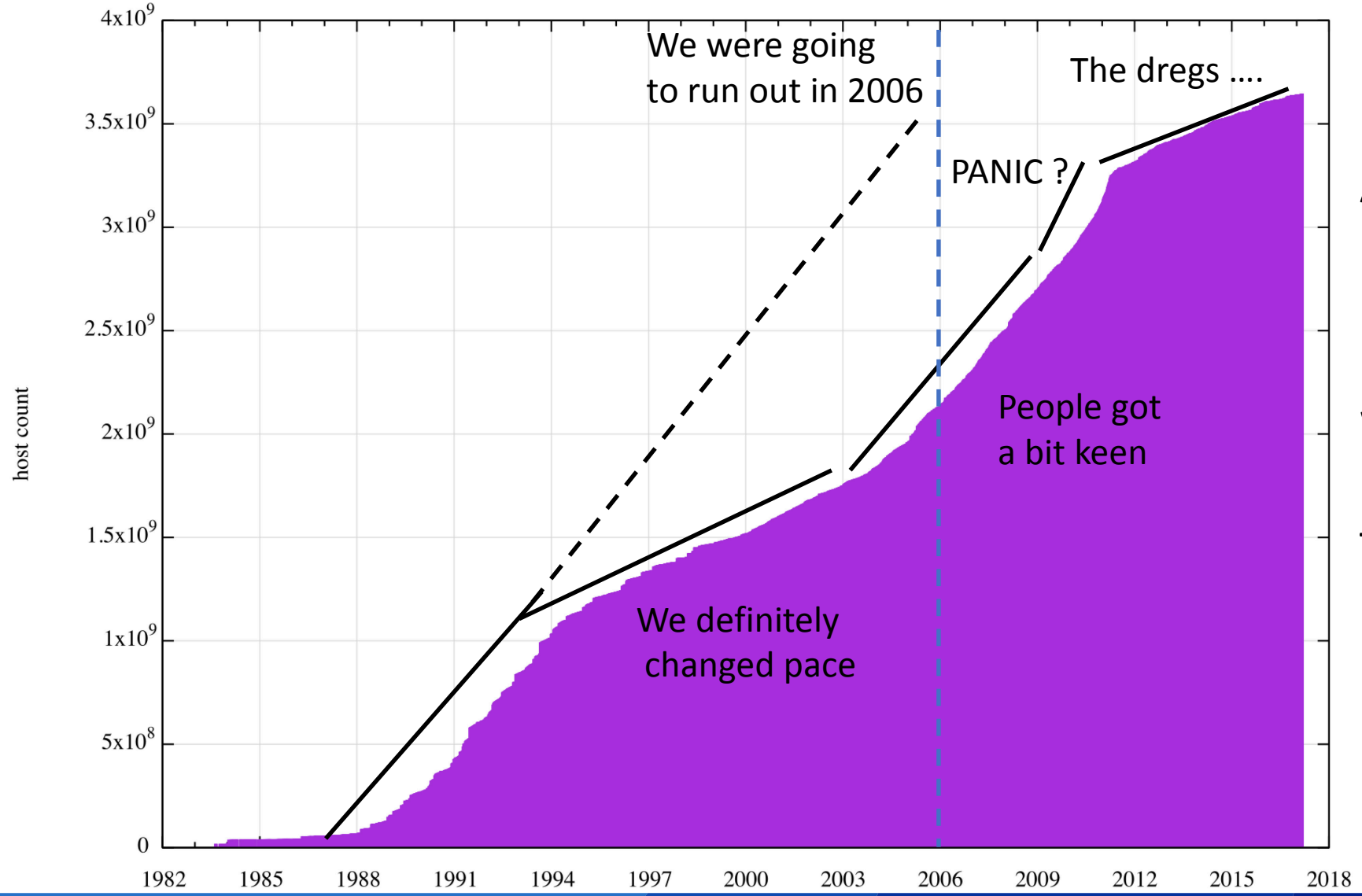
ggm@apnic.net

With thanks to {martin,olafur}@cloudflare.com

Overview

- Background “we knew this was coming”
  - Final /8 process
  - Tainted blocks testing: AARNet and Google
  - Failed APNIC policy attempt: reserved to labs
- The cloudflare proposal
- How is it going (cloudflare material)
- Where to from here?

Address policy in five lines (and one curve)



Address policy
Probably bought
10 years.

We didn't succeed
In using the 10 years
To deploy IPv6.

We knew this was coming..

- BGP the movie
- Rundown prediction models from APNIC & others
  - 2008-2010 emerging final /8 policy proposals in all RIR.
- Forseeable sometime between 2010 and 2012, we'd need to plan for end of supply in IPv4 from IANA
- "Awkward" /8 assignments: Leo Vegoda (IANA) September 2007
  - Mentions 1.0.0.0/8 5.0.0.0/8 and 42.0.0.0/8 (APNIC got 1 and 42. Ripe got 5)
  - https://www.cisco.com/c/en/us/about/press/internet-protocol-journal/back-issues/table-contents-37/103-awkward.html
- https://www.icann.org/news/blog/selecting-which-8-to-allocate-to-an-rir
  - Avoid pain for AfriNIC and LacNIC: give them preferential access to untainted blocks
  - Random assignment of tainted blocks to APNIC, ARIN & RIPE NCC (otherwise in line with address policy, rundown planning)

Lets talk about 1.0.0.0/8 specifically

- 1.0.0.0/8 (1/8) reserved by IANA, since 1981.
  - used unofficially as example addresses, default configuration parameters or pseudo-private address space.
  - 1.1.1.0/24 is used by Intel bladecenters internally for communication with its blades.
  - Used by MacDonalds for free WiFi, Swisshotels and others, in documentation for configuration examples of products
    - And see later for what CloudFlare have uncovered for continuing uses
- Status changed in IANA from ”reserved” to “unallocated” in 2008
  - Predictions about runout were within foreseeable timeframe

1.0.0.0/8 Assigned to APNIC January 2010

- Tested by APNIC across 2010 with a range of partners
  - RIPE NCC announcing more specifics
    - see 50 mbit/second, floods IX links
    - https://labs.ripe.net/Members/franz/content-pollution-18
  - MERIT peak burst 860 Mbps
    - Audio data, traffic levels unlike other tested netblocks in the ‘awkward’ set
    - https://www.merit.edu/wp-content/uploads/2018/01/1.0.0.08.pdf
  - Google/YouTube
    - traffic levels of around 150 Mbps mainly comprising UDP (cannot easily be constrained) 80Mbps in 1.1.0.0/16
    - Peaks over 800Mbps
  - AARNet
- Five sub-ranges identified for hold-back:
  - 1.0.0.0/24 10+ Mbps
  - 1.1.1.0/24 80+ Mbps
  - 1.2.3.0/24 10+ Mbps
  - 1.4.0.0/24 10+ Mbps
  - 1.10.10.0/24 3+ Mbps
    - http://www.potaroo.net/studies/1slash8/1slash8.pdf

Final /8 policy implications

- IANA exhausted its IPv4 free pool (3 February 2011)
- From apnic's web pages:
  - https://www.apnic.net/community/ipv4-exhaustion/graphical-information/

    On 15 April 2011, the APNIC pool reached the last /8 of available IPv4 addresses, triggering the Final /8 policy.

    :

    :

    Quarantined blocks will be released to the Available pool when their routability and usability problems are minimized.
- Pre-exhaustion testing implemented by labs, 2009-2011
- Tainted ranges identified and held back
  - Too much traffic inbound, to make acceptable for routine distribution

What to do with 1.1.1.0/24 and 1.0.0.0/24

- Many competing requests: CNNIC, Microsoft
  - Rejected at OPM on the floor (or failed to reach consensus on the ML)
- 26 January 2014 Labs (GIH) proposes retain for research
  - 30 April 2014 endorsed by APNIC EC.
  - 7 May 2014 implemented as policy. Blocks marked in WHOIS to Labs.
- Labs holds the block, continues assessments with Google, AARNet.
  - Provide strong signal of “background traffic” levels. Remain infeasibly tainted for routine use
- Enter Cloudflare..

August 2017 request from Cloudflare

- Cloudflare contacted us asking to enter a research relationship with these ranges
  - Can they explore sinking this load in their global anycast, and offering public DNS?
- No implied right to assignment
  - Blocks remain in the control of Labs, under address policy
  - Future/continuing use has to be understood to be bound by address policy
- Strong privacy drivers for service (from Cloudflare)
  - No sharing of endpoint IP, queried domains, No data taken from core DNS system
  - Simple APIs for bulk data information: volumes, origin-AS type information
- March 2018, Registry records updated, LOA published
- April 1 2018 Cloudflare announce services in BGP

Cloudflare slides

Announced April 1st 2018

Our mission: to help build a better Internet.

We use 1.1.1.1 and 1.0.0.1 (easy to remember) for our resolver.

Addresses provided to Cloudflare by APNIC for both joint research and this service.

We focused on privacy!

We knew we would spend a lot of time cleaning up the global Internet to make 1.1.1.1 work!

https://blog.cloudflare.com/announcing-1111/
https://blog.cloudflare.com/dns-resolver-1-1-1-1/

CLOUDFLARE

1.1.1.1

DNS resolver, 1.1.1.1, is served by Cloudflare's Global Anycast Network.

APNIC Labs and Cloudflare

APNIC Labs enters into a research agreement with Cloudflare

By Geoff Huston on 2 Apr 2018

Category: Tech matters

Tags: DNS, Research

APNIC Labs is partnering with Cloudflare for a joint research project relating to the operation of the DNS.

I'd like to explain our motivation in entering into this research project, explain what we hope to be able to achieve with this work, and describe briefly how we intend to handle the data that will be generated from this research activity.

The joint research project involves the operation of an open public DNS resolution service using IPv4 address prefixes that the APNIC Address Policy Special Interest Group (SIG) has set aside for research purposes. This project will provide APNIC Labs with unique opportunity to gain valuable insight into the query behaviour of the DNS in today's Internet and will allow us to further our existing research activities in looking at the DNS.

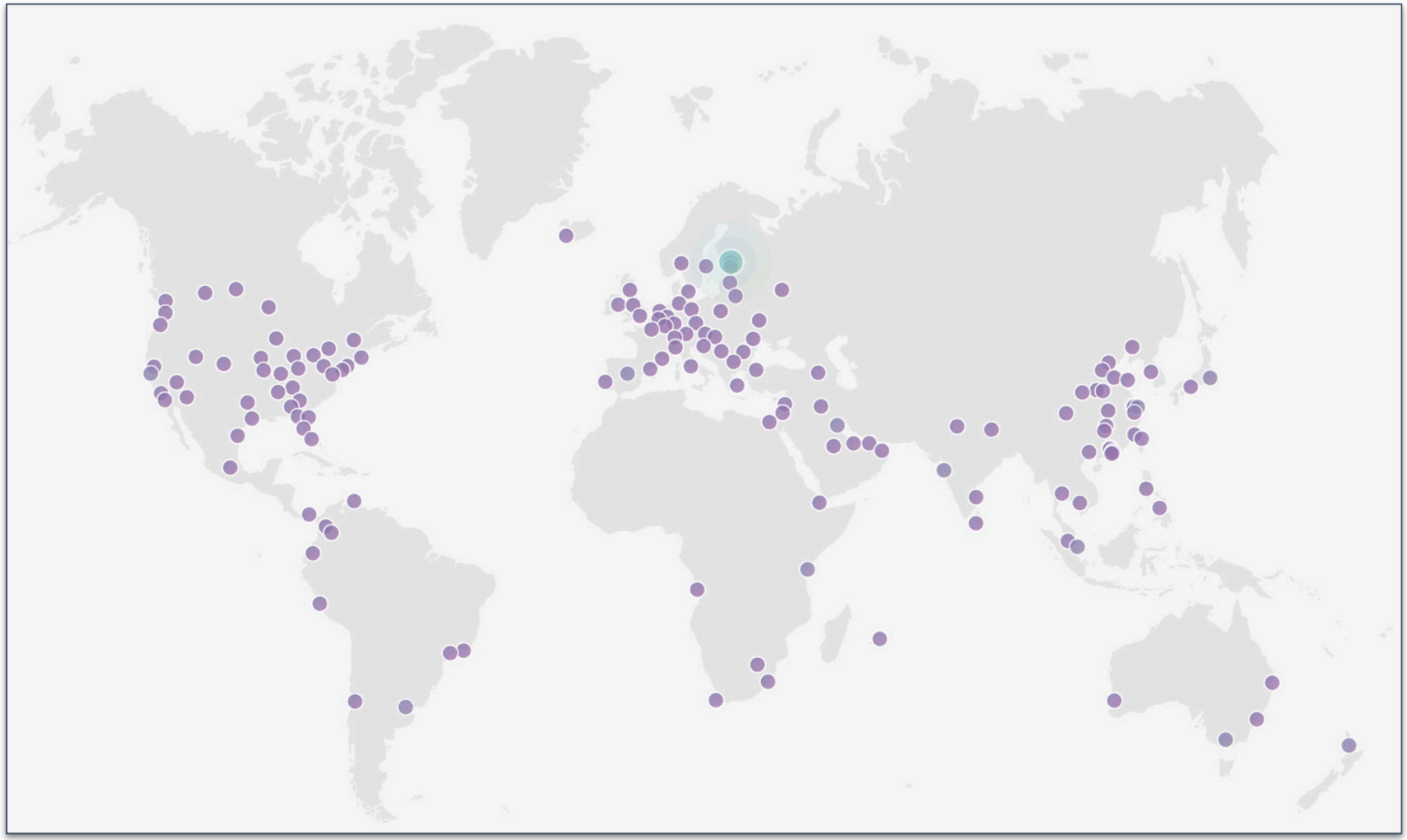
https://blog.apnic.net/2018/04/02/apnic-labs-enters-into-a-research-agreement-with-cloudflare/

CLOUDFLARE

1.1.1.1

APNIC is allocated 1.0.0.0/8 by IANA in January 2010

The Cloudflare network (DNS, DDoS, CDN, WAF, more)



151+

Data centers globally

151+

DNS resolver locations

151+

DNS authoritative locations

1.1.1.1 design goals

DNS and privacy!

DNS itself is a 35-year-old protocol (and it's showing its age). It was never designed with privacy or security in mind.

DNS inherently is unencrypted so it leaks data to anyone who's monitoring your network connection.

We focused on privacy:

- Query Minimization RFC7816
- Aggressive negative answers RFC8198
- No Client Subnet on queries

- DNS-over-TLS (Transport Layer Security) RFC7858
- DNS-over-HTTPS protocol DoH (draft-ietf-doh-dns-over-https)

CLOUDFLARE

1.1.1.1

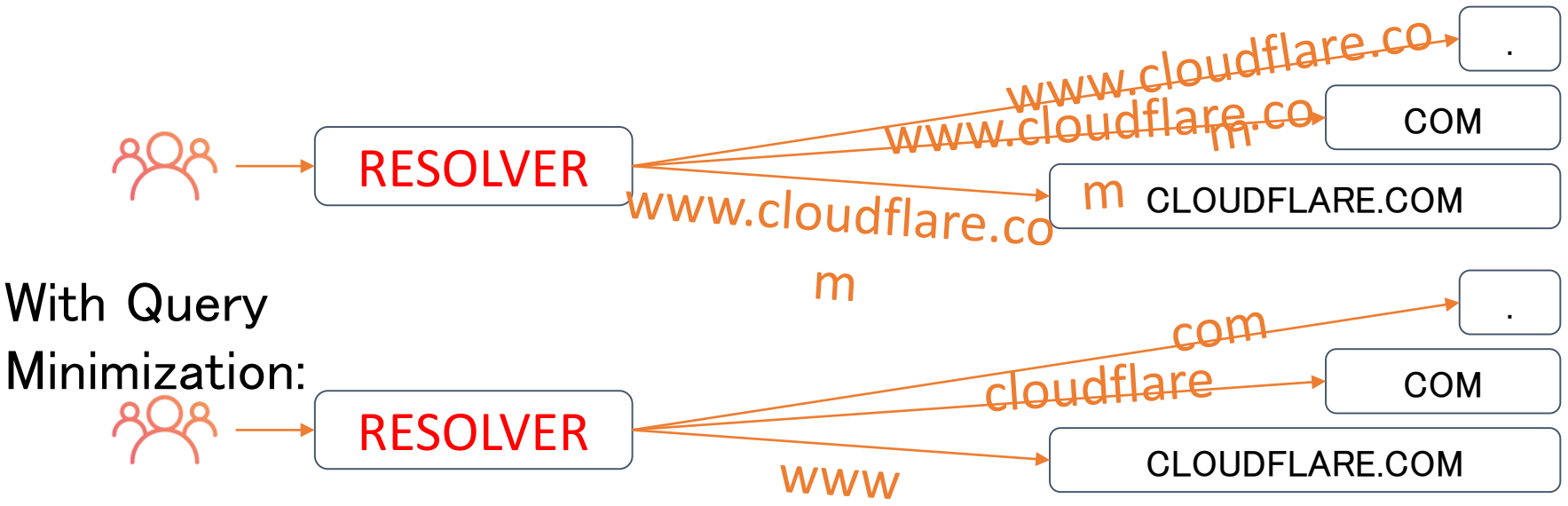
In 2014, we decided to enable https encryption for free for all our customers (we doubled the size of the encrypted web).

In 2017, we made DDoS mitigation free & unmetered across all our plans.

DNS Query Minimization

- DNS is chatty, very chatty!

- Resolver can reduce the information leaked to intermediary DNS servers
  - The root, TLDs, and secondary zones
- Resolver only sends just enough of the name for the authority to tell the resolver where to ask the next question.



CLOUDFLARE

1.1.1.1

QNAME contains too much information.

DNS Aggressive Negative Answer

- Fewer lookups to authorities (in particular the root zone)
- Use the existing resolvers negative cache
  - Negative (or non-existent) information kept around for a period of time

- For zones signed with DNSSEC with the NSEC records in cache:
  - Resolver can figure out if the requested name does NOT exist without doing any further queries
  - If you type wwwwwww dot something and then www dot something, the second query could well be answered with a very quick “no” (NXDOMAIN in the DNS world)

- Aggressive negative caching works only with DNSSEC signed zones, which includes both the root and ~1,400 out of 1,544 TLDs

1.1.1.1

QNAME contains too much information.

CLOUDFLARE

Client Subnet == Bad privacy

Client Subnet: RFC7871/Experimental

- Used for traffic engineering when queries come from open resolvers or large resolver clusters
  - addr/netmask ⇒ fine grain “location” /24 commonly used
  - Bad for resolvers as it kills cache hit ratio
  - Resolver cache implementations got more complex
- Suggestions to use it to track devices behind a NAT

Not using ECS degrades performance in some cases

Fine grain steering vs course steering

Where should traffic steering actually happen?

- DNS
- Applications via referrals ?

What is acceptable scope for NetMask ?

CLOUDFLARE

1.1.1.1

CS option frequently included on all queries ⇒
Massive data leak

How to find the right balance?

DNS-over-TLS / DNS-over-HTTPS

TLS (Transport Layer Security) is the basis of https encryption.

- DNS-over-TLS (RFC7858) is simply a DNS request(s) wrapped by TLS.
- DNS-over-HTTPS (draft-ietf-doh-dns-over-http) is DNS queries via an HTTPS request. **

Resolver, 1.1.1.1 now provides both - at scale!

- Mozilla Trusted Recursive Resolver
  - Cloudflare listed

** https://hacks.mozilla.org/2018/05/a-cartoon-intro-to-dns-over-https/
https://daniel.haxx.se/blog/2018/06/03/inside-firefoxs-doh-engine/

CLOUDFLARE

1.1.1.1

DNSSEC ensures integrity of data between resolver and authoritative server, it doesn't protect privacy of that data!

Specifically, DNSSEC doesn't protect the privacy of the "last mile".

Data Policy

- We don’t store client IP addresses never, ever!
- We only use query names for things that improve DNS resolver performance.
- After obfuscation, APNIC research gets access to data (under our joint agreement).

- Cloudflare never stores any information in logs that identifies end user.
    - All log records are deleted within 24 hours.
- We will continue to abide by our privacy policy and ensure that no user data is sold to advertisers or used to target consumers.

1.1.1.1

All log records deleted within 24 hours

CLOUDFLARE

DNS resolver addresses

IPv4 & IPv6

1.1.1.1

1.0.0.1

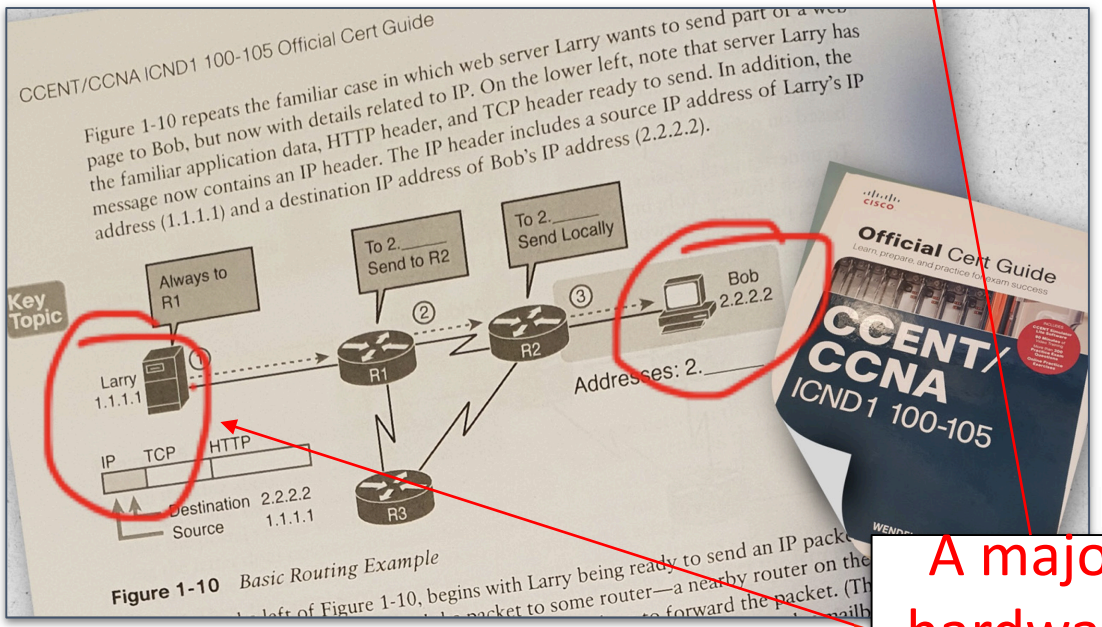
2606:4700:4700::1111

2606:4700:4700::1001

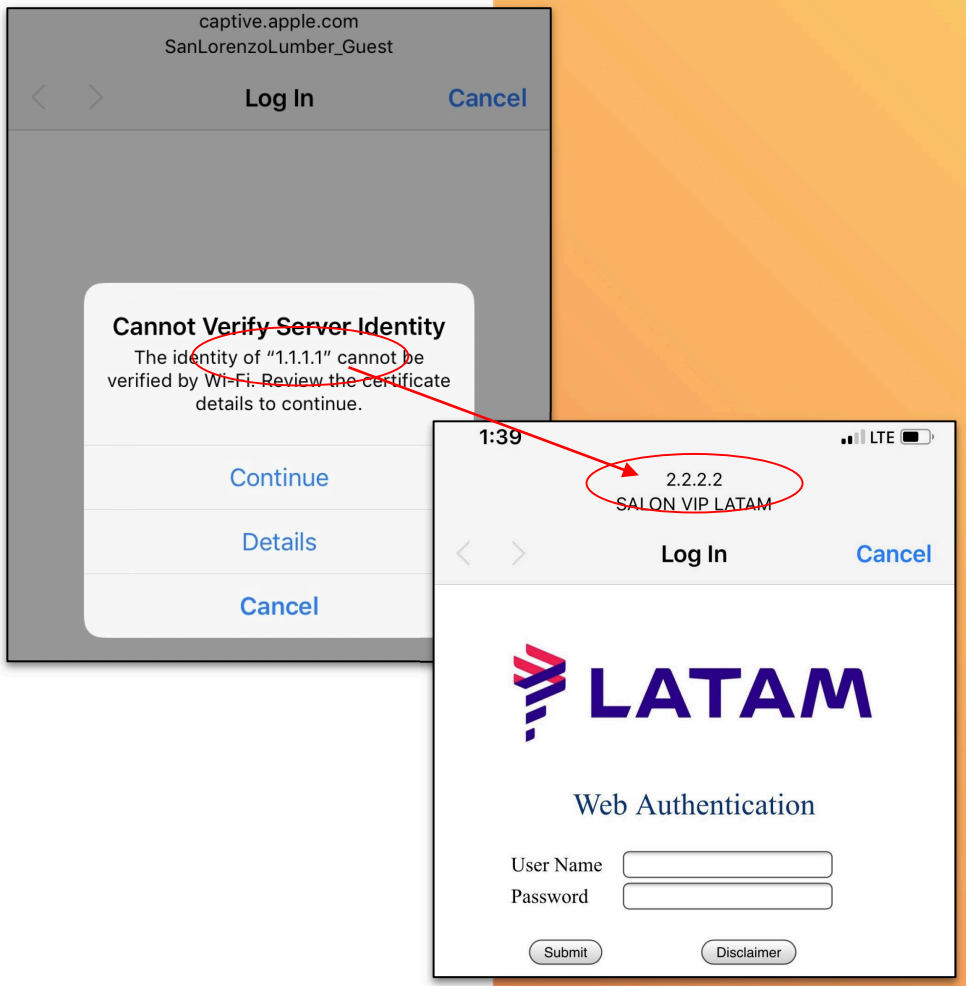
1.1.1.1 polluted space

Step 32 In the IP Address text box, enter the IP address of the controller's virtual interface.

You should enter a fictitious, unassigned IP address such as 1.1.1.1.



A major hardware vendor



1.1.1.1

Polluted for many many years

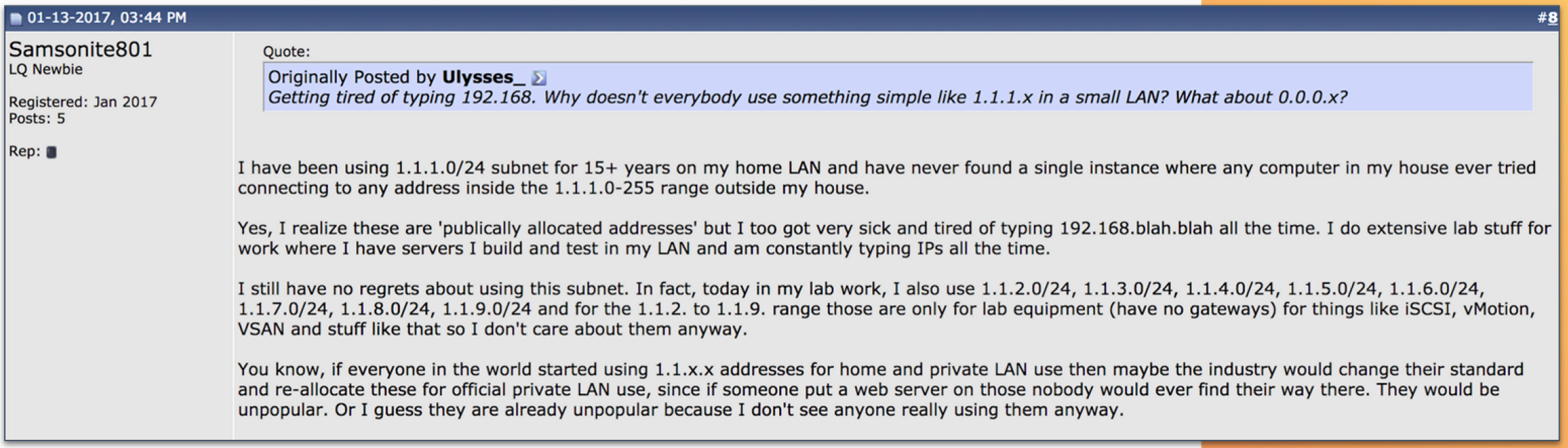
CLOUDFLARE

1.1.1.1 polluted space

1.1.1.1

Hard to explain “assigned” vs “private”

Sadly, user “Samsonite801” will never be able to use 1.1.1.1 DNS resolver!



01-13-2017, 03:44 PM #8

Samsonite801
LQ Newbie

Registered: Jan 2017
Posts: 5

Rep:

Quote:

Originally Posted by **Ulysses_**
Getting tired of typing 192.168. Why doesn't everybody use something simple like 1.1.1.x in a small LAN? What about 0.0.0.x?

I have been using 1.1.1.0/24 subnet for 15+ years on my home LAN and have never found a single instance where any computer in my house ever tried connecting to any address inside the 1.1.1.0-255 range outside my house.

Yes, I realize these are 'publically allocated addresses' but I too got very sick and tired of typing 192.168.blah.blah all the time. I do extensive lab stuff for work where I have servers I build and test in my LAN and am constantly typing IPs all the time.

I still have no regrets about using this subnet. In fact, today in my lab work, I also use 1.1.2.0/24, 1.1.3.0/24, 1.1.4.0/24, 1.1.5.0/24, 1.1.6.0/24, 1.1.7.0/24, 1.1.8.0/24, 1.1.9.0/24 and for the 1.1.2. to 1.1.9. range those are only for lab equipment (have no gateways) for things like iSCSI, vMotion, VSAN and stuff like that so I don't care about them anyway.

You know, if everyone in the world started using 1.1.x.x addresses for home and private LAN use then maybe the industry would change their standard and re-allocate these for official private LAN use, since if someone put a web server on those nobody would ever find their way there. They would be unpopular. Or I guess they are already unpopular because I don't see anyone really using them anyway.

https://www.linuxquestions.org/questions/linux-networking-3/why-doesn%27t-everyone-use-1-1-1-x-or-1-1-x-x-or-1-x-x-x-addresses-in-their-lans-4175563056/

CLOUDFLARE

1.1.1.1 polluted space (the edge)

Many CPE routers use 1.1.1.1 for captive portals or configuration screens

- Pace (Arris) 5268
- D-Link DMG-6661
- Technicolor C2100T
- Calix GigaCenter ---- fixed 2018/Jun/12 thanks to a USER
- Nomadix (model(s) unknown)
- Xerox Phaser MFP

Deployed in the millions globally

1.1.1.1

Millions of CPE boxes globally

CLOUDFLARE

1.1.1.1 polluted space (backbones)

Many backbones seem to have 1.1.1.1 backholed or used - for no real reason

We committed to fixing this by using our measurements to track down, contact and correct these inconsistencies. Here's a partial list of successfully cleaned backbones!

- Airtel, BHTelecom, Beirut-IX, Comcast, Fastweb, ITC, Kazakhtelecom, LG Telecom, Level(3), Liquid Telecom, MTN, Omantel, Rostelecom, SFR, SKBB, Sonatel, STC, Tata, Telecom Italia, Telenor, Telus, Turk Telekom, Turkcell, Voo, XS4ALL, Ziggo
- Many more ...

Thank you backbones. You have helped the Internet improve.

CLOUDFLARE

1.1.1.1

Why do backbones use this route?

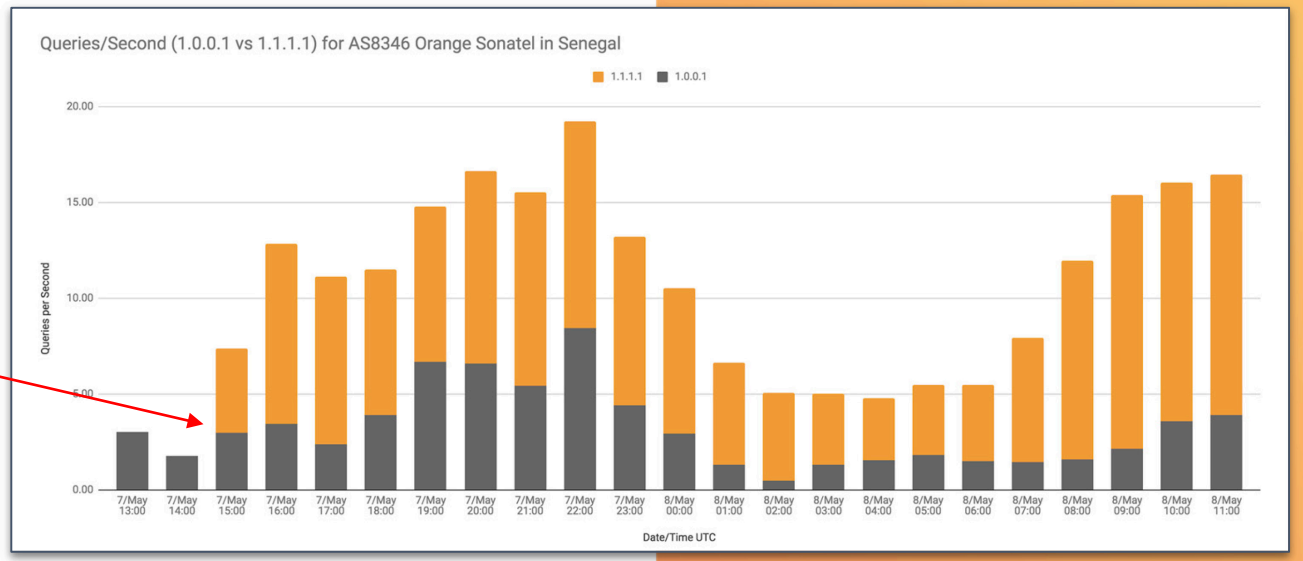
Good question!

1.1.1.1 fixed in Senegal

- 1.1.1.1 (1.1.1.0/30) was in use internally within Sonatel
  - This isn't unusual - (see previous slides)
  - Prevents end-users from accessing resolver at 1.1.1.1
  - However, 1.0.0.1 is available - hence resolver always worked

- This is repeated in many countries and telcos

Fixed!



1.1.1.1

Fixing 1.1.1.1, one network at a time!

CLOUDFLARE

Measuring availability

- Thanks to the RIPE Atlas probes and thousands of tests
  - Tested ISPs globally for access to 1.1.1.1 (and 1.0.0.1)
  - Sent many emails to many NOCs **

Time (UTC)	RTT		Hops	Success	
2018-03-28 11:43	7.504		11	✗	
2018-03-28 11:43	6.292		11	✗	
2018-03-28 11:43	6.260		11	✗	
2018-03-28 11:43	8.558		11	✗	
2018-03-28 11:43	7.308		11	✗	
2018-03-28 11:43	3.412		11	✗	
2018-03-28 11:43	33.123		11	✗	
2018-03-28 11:43	1.879		1	✓	
2018-03-28 11:43	21.928		7	✓	
2018-03-28 11:43	11.641		8	✗	
2018-03-28 11:43	26.318		4	✓	

Null-routes

CPE installed in ISP

…

Suddenly an open FTP server

CLOUDFLARE

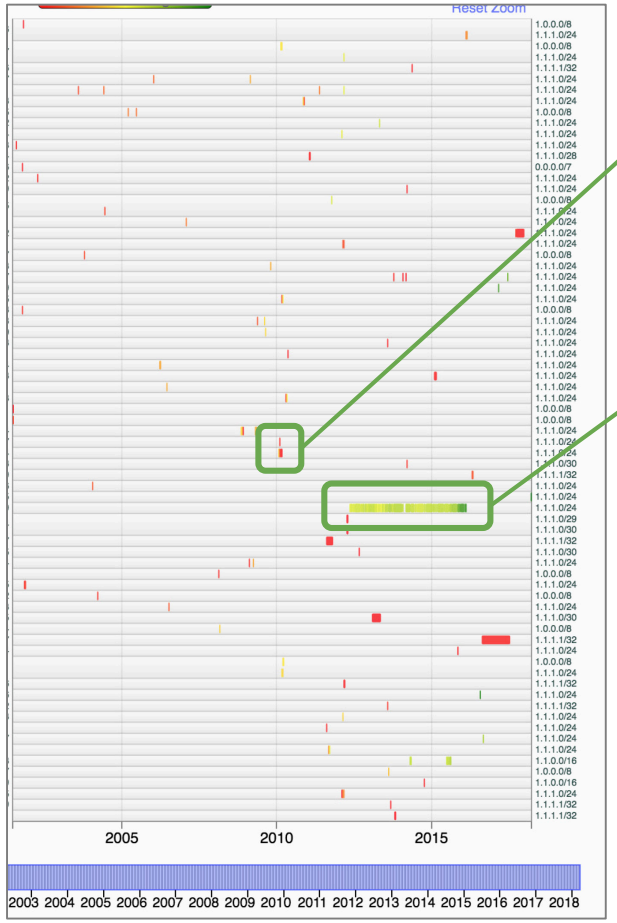
** https://blog.cloudflare.com/fixing-reachability-to-1-1-1-1-globally/

1.1.1.1

RIPE Atlas to the rescue!

1.0.0.0/24 & 1.1.1.0/24 background noise

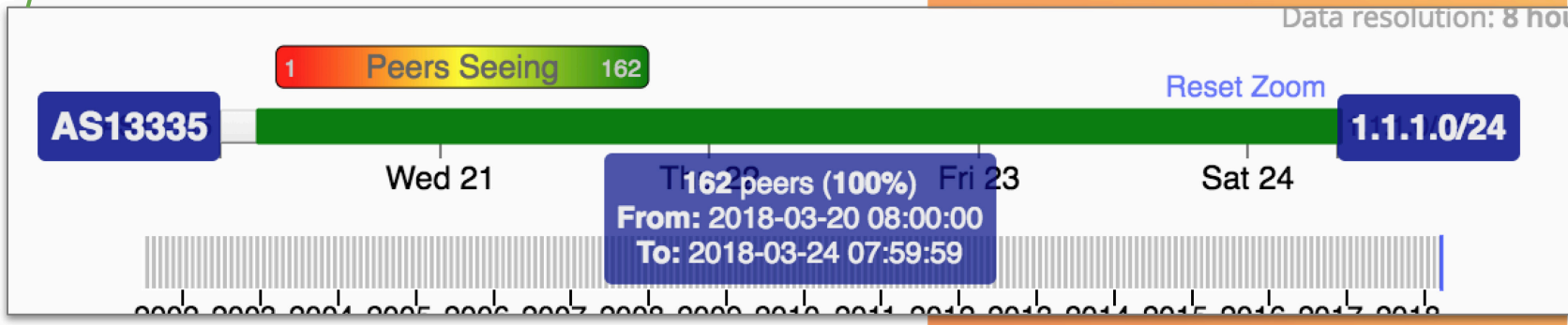
1.1.1.0/24 routing history



RIPE, Merit
https://labs.ripe.net/Members/franz/content-pollution-18
– Franz Schwarzinger
http://www.potaroo.net/studies/1slash8/1slash8.html
– Geoff Huston

Google, YouTube

AS13335 Cloudflare



CLOUDFLARE

1.1.1.1

10+ Gbps of noise!

1.1.1.0/24 background traffic

1.1.1.1

10+ Gbps of noise!

1.0.0.0/24 gets about 1%

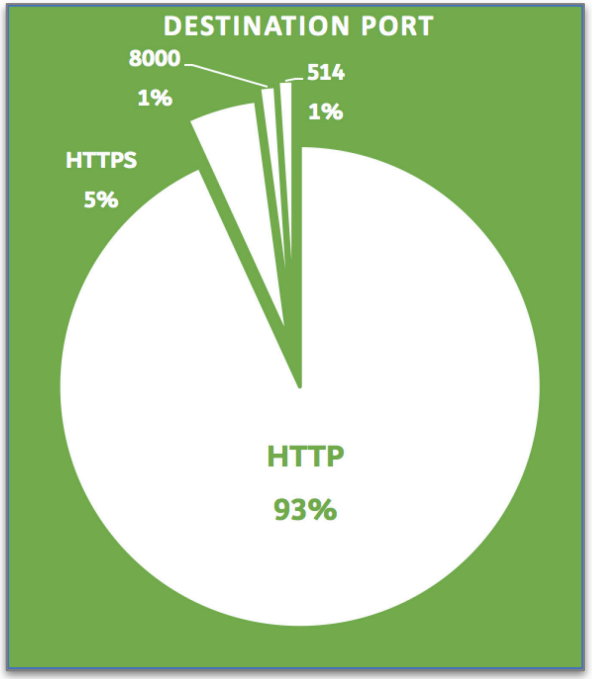
- Previous studies:
  - **2010:** Greater than 100 Mbps on 1.1.1.0/24
  - **2014:** 100 Mbps → 1 Gbps on 1.0.0.0/8 **
- Cloudflare routing:
  - **2018:** 8 Gbps → 13 Gbps (with 1 Gbps solely on 1.1.1.1)

** https://conference.apnic.net/data/37/2014-02-27-prop-109_1393397866.pdf
– Geoff Huston

CLOUDFLARE

1.1.1.0/24 background traffic

- TCP traffic (mostly HTTP proxy, services).
  - Ports 80, 443, 8000, 8080, 8090, 8765
- UDP traffic (some DNS, syslogs).
  - Ports 53, 514, 8000, 80, 8090
- TP-Link DNS 1.0.0.19 **



TP-Link routers send DNS queries to 1.0.0.19. What is that?

4

I've got a problem with TP-Link soho routers. The DNS forwarder of those routers tends to ignore the DNS servers obtained by DHCP and instead tries sending all DNS requests to this strange IP: 1.0.0.19? That IP doesn't respond.

Has anyone else seen that happen?

domain-name-system

** https://serverfault.com/questions/365613/tp-link-routers-send-dns-queries-to-1-0-0-19-what-is-that/365630

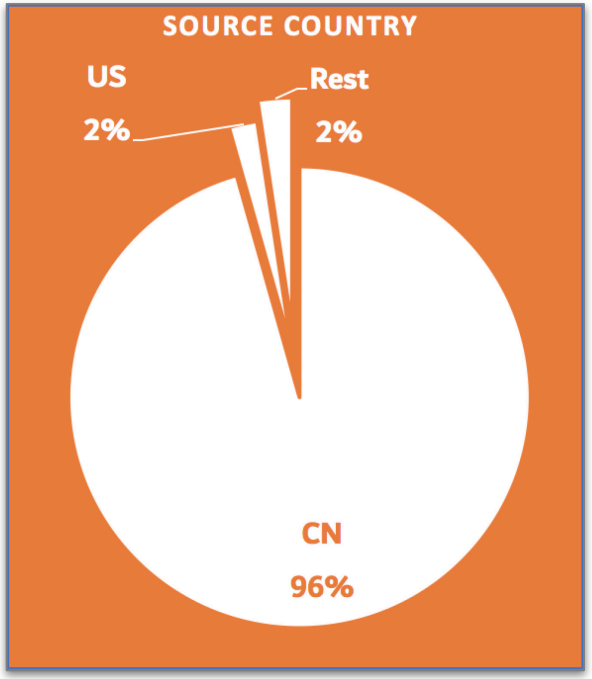
CLOUDFLARE

1.1.1.1

10+ Gbps of noise!

1.1.1.0/24 background traffic

- Traffic source
  - Mostly China
  - US
  - countries in Asia
  - some Europe



1.1.1.1

10+ Gbps of noise!

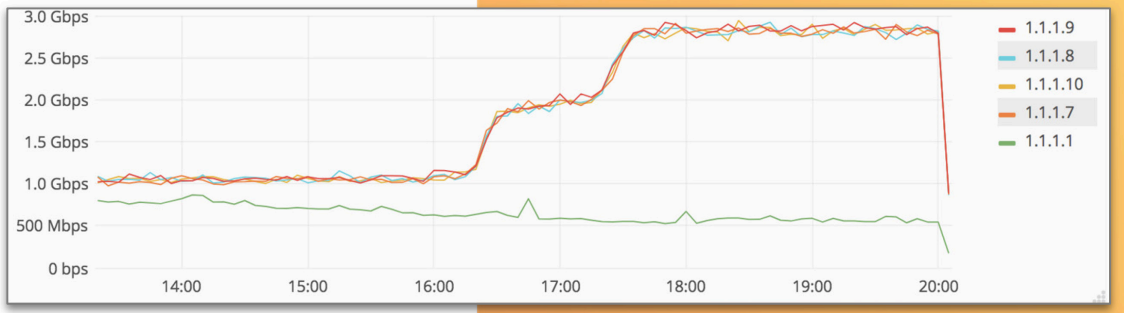
CLOUDFLARE

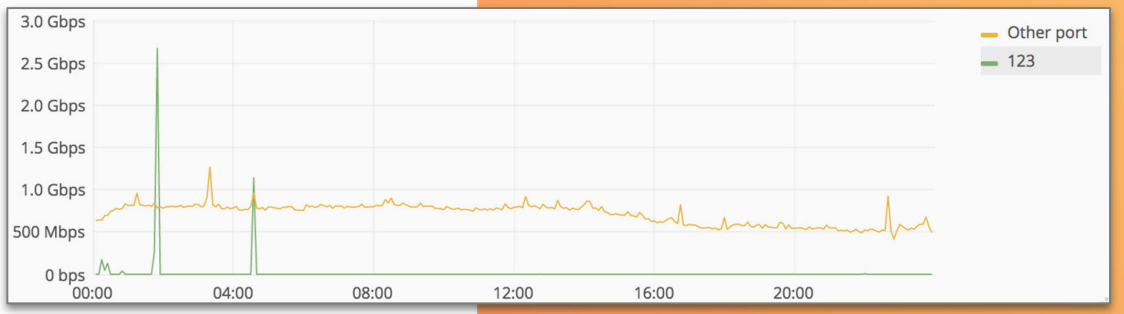
1.1.1.0/24 bursts and patterns

1.1.1.1

10+ Gbps of noise!

- Two increases:
  - 5 Gbps → 8 Gbps between 16:00 → 17:15 UTC
  - 8 Gbps → 12.5 Gbps between 17:15 → 23:00 UTC
  - Mostly on 1.1.1.7, 1.1.1.8, 1.1.1.9, and 1.1.1.10
    - Destination port 80
    - Increase from China
    - No particular difference on source IP/net
- Short bursts:
  - Only on 1.1.1.1 between 01:00 → 02:00 UTC for a few minutes
  - 1 Gbps → 10 Gbps
  - UDP traffic source port 123 (NTP) and port 11211 (memcached)
    - Misconfigured network devices?





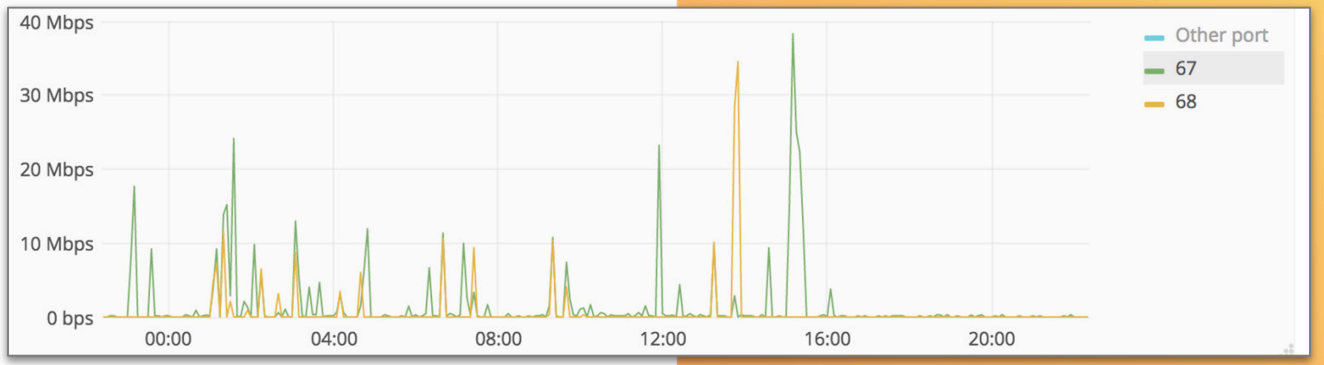
CLOUDFLARE

1.1.1.0/24 bursts and patterns

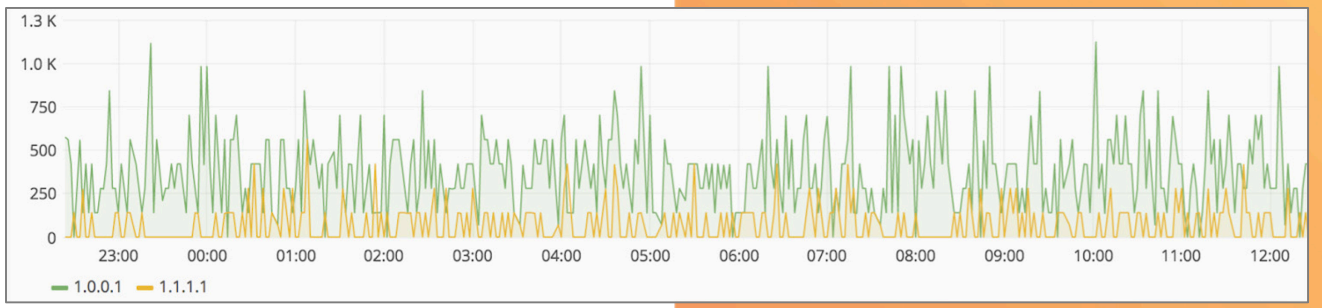
1.1.1.1

- Also DHCP spikes from Macau
  - Bursts to 40 Mbps

10+ Gbps of noise!



- How many packets per second on UDP 53 (before launching)



CLOUDFLARE

1.1.1.0/24 what changed?

- Presentation from 10 years ago at NANOG49 **
  - *“iperf traffic to 1.2.3.4 is roughly 10 Mbps of traffic from less than a 100 unique sources”*
- 2018: we still see iperf traffic (port 5000/5001)
  - Around 10-20 times the traffic

We estimate legitimate traffic to be around **7–13%**

** https://www.nanog.org/meetings/nanog49/presentations/Monday/karir-1slash8.pdf
Merit, APNIC, University of Michigan

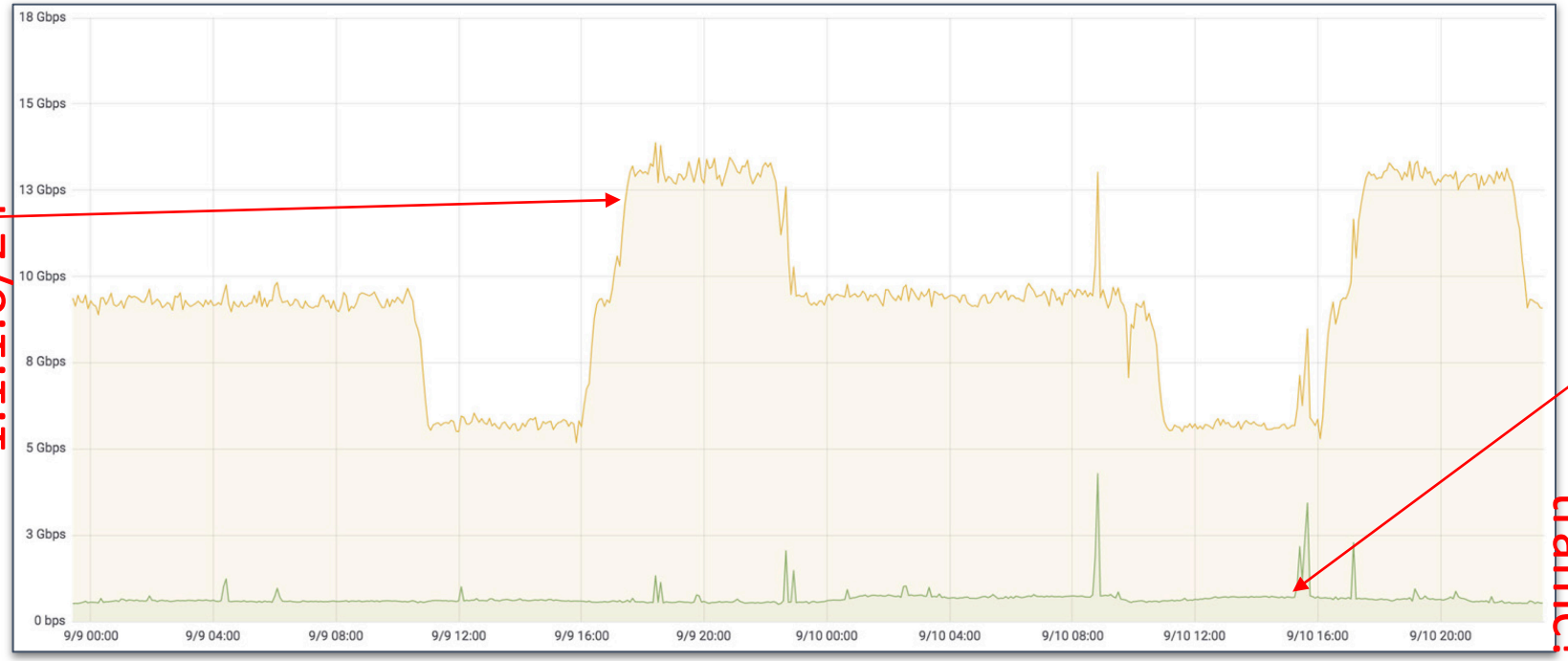
1.1.1.1

10+ Gbps of noise!

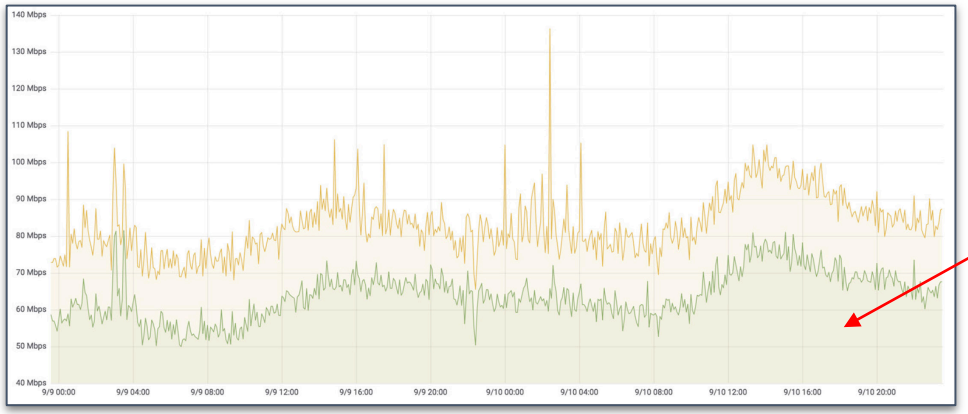
CLOUDFLARE

1.0.0.0/24, 1.1.1.0/24 traffic

All traffic to 1.1.1.0/24



1.1.1.1 resolver traffic!



1.0.0.1 resolver traffic!

CLOUDFLARE

1.1.1.1

10+ Gbps of noise!

1.1.1.1 @ ~ 600 Mbps
1.0.0.1 @ ~ 70 Mbps

1.1.1.0/24 noise somewhat down
1.0.0.0/24 noise significantly down

1.0.0.0/24, 1.1.1.0/24 traffic

1.1.1.1 traffic!

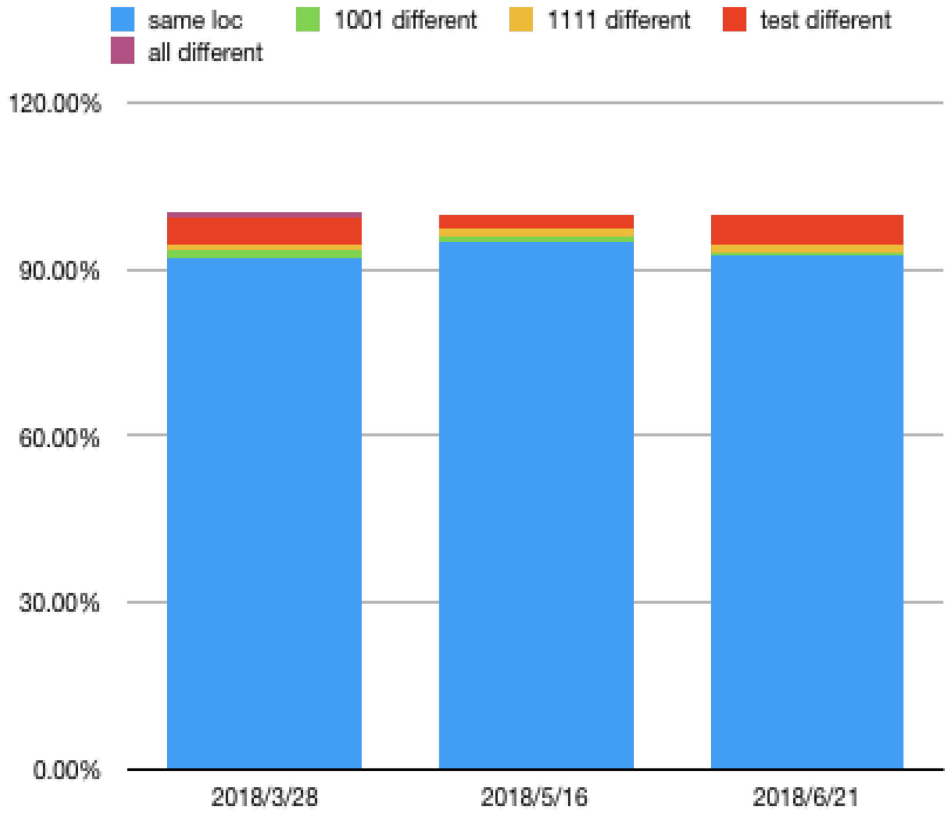
CLOUDFLARE

1.1.1.1

1.1.1.0/24 noise

Routing

Traffic goes where ?



Not all go to same location

Date	1.0.0.1	1.1.1.1	Test	#
Mar/28	8.3%	14.7%	4.8%	16.7%
May/16	0.4%	3.0%	0.2%	3.4%
Jun/21	1.2%	4.2%	1.5%	5.0%

Reachability issues persist

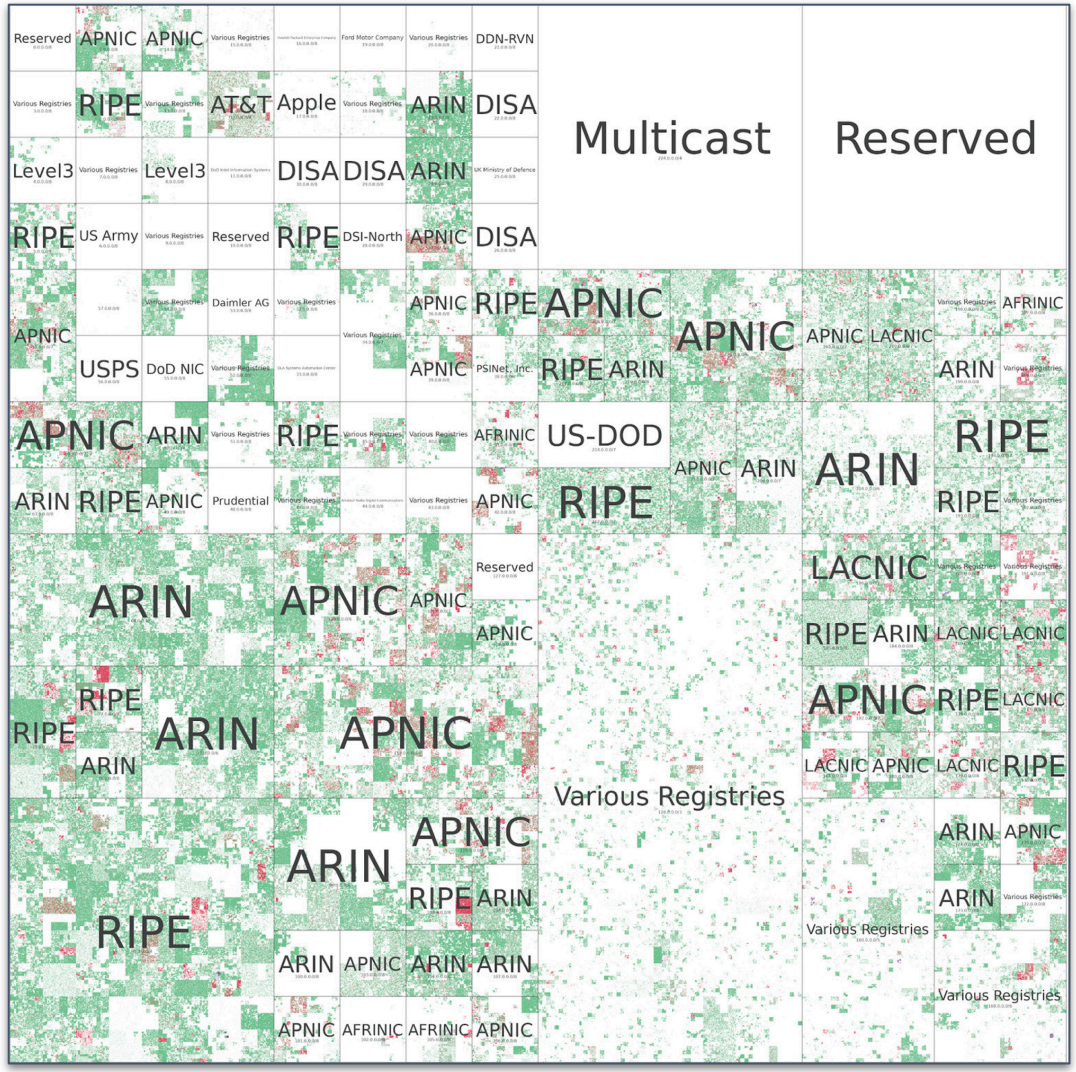
CLOUDFLARE

1.1.1.1

Measured from Ripe Atlas probes

Old Tunnels never die

Measuring availability (via pings)



1.1.1.1

Resolver reachability

Green - All working
Red = 1.1.1.1 fails
Pink = 1.0.0.1 fails
Purple = both fail

Early August/2018

Hilbert curves are cool!

CLOUDFLARE

Captive Portals are the worst

Debug Information

Connected to 1.1.1.1	No
Using DNS over HTTPS (DoH)	No
Using DNS over TLS (DoT)	No
AS Name	Massachusetts Institute of Technology
AS Number	3
Cloudflare Data Center	BOS

Connectivity to Resolver IP Addresses

1.1.1.1	No
1.0.0.1	Yes
2606:4700:4700::1111	No
2606:4700:4700::1001	No

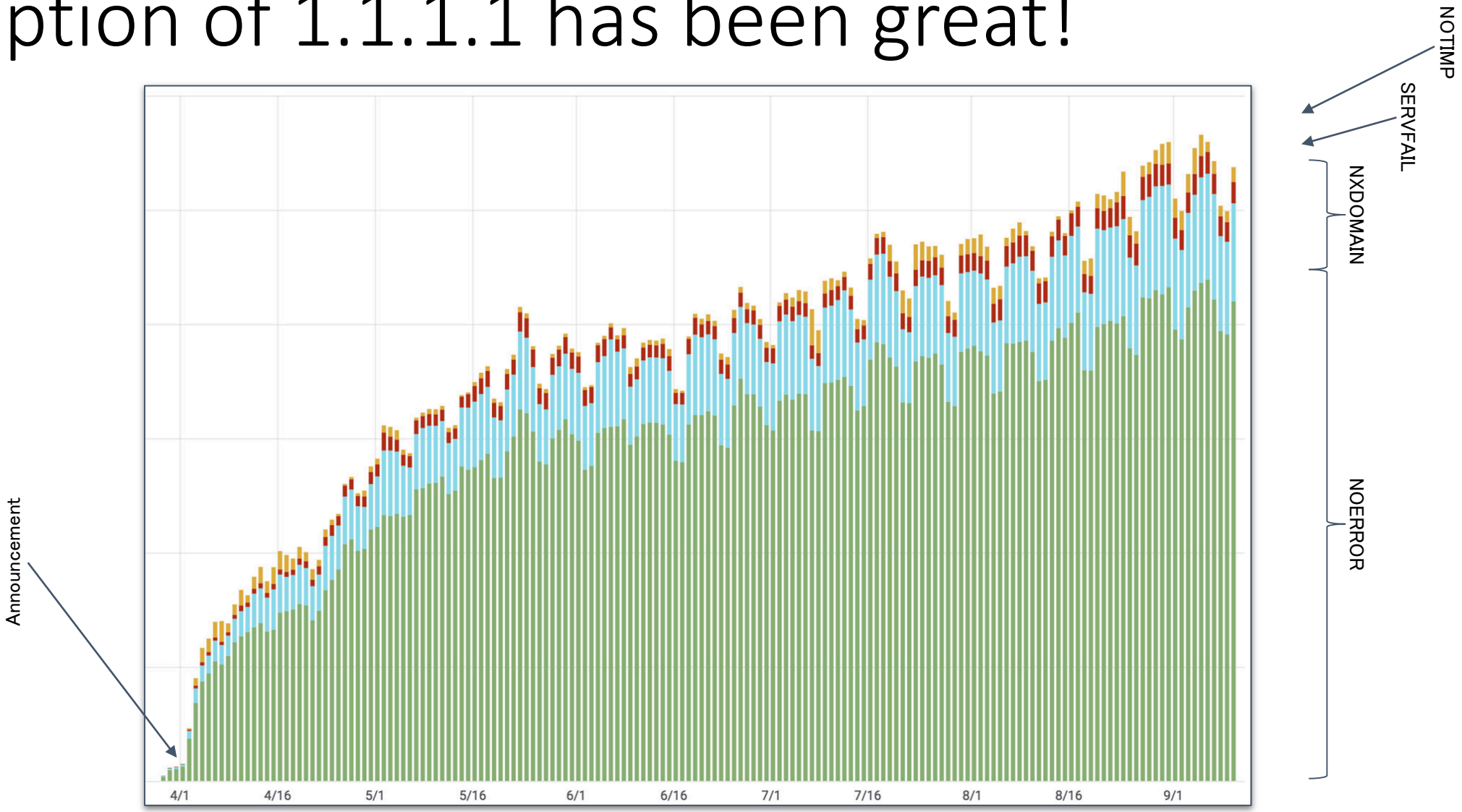
CLOUDFLARE

1.1.1.1

MIT Guest network at 22/6/2018 10:14

Adoption

Adoption of 1.1.1.1 has been great!



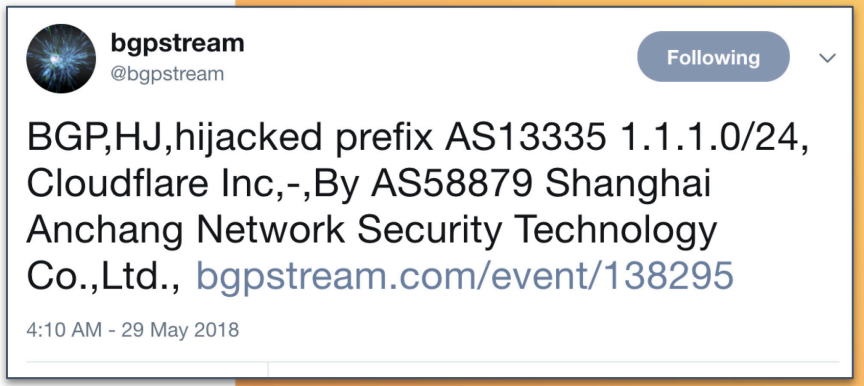
About route leaks

1.1.1.0/24 leaks happen

- The heavy use of 1.1.1.1 in networks (running BGP) trigger route leaks
- Cloudflare has a signed RPKI ROA for both 1.0.0.0/24 & 1.1.1.0/24
  - RPKI signed - but doesn't (yet) stop route leaks
- The 29 May 2018 leak was ~60 seconds in length
  - It lasted longer on twitter
- This must stop; not just for this route, but on all routes!

1.1.1.1

Route leaks need to stop!




```
Prefix:          1.1.1.0/24
Country code:    AU
Origin AS:       13335
Origin AS Name:  Cloudflare Inc
RPKI status:     ROA validation successful
```

CLOUDFLARE

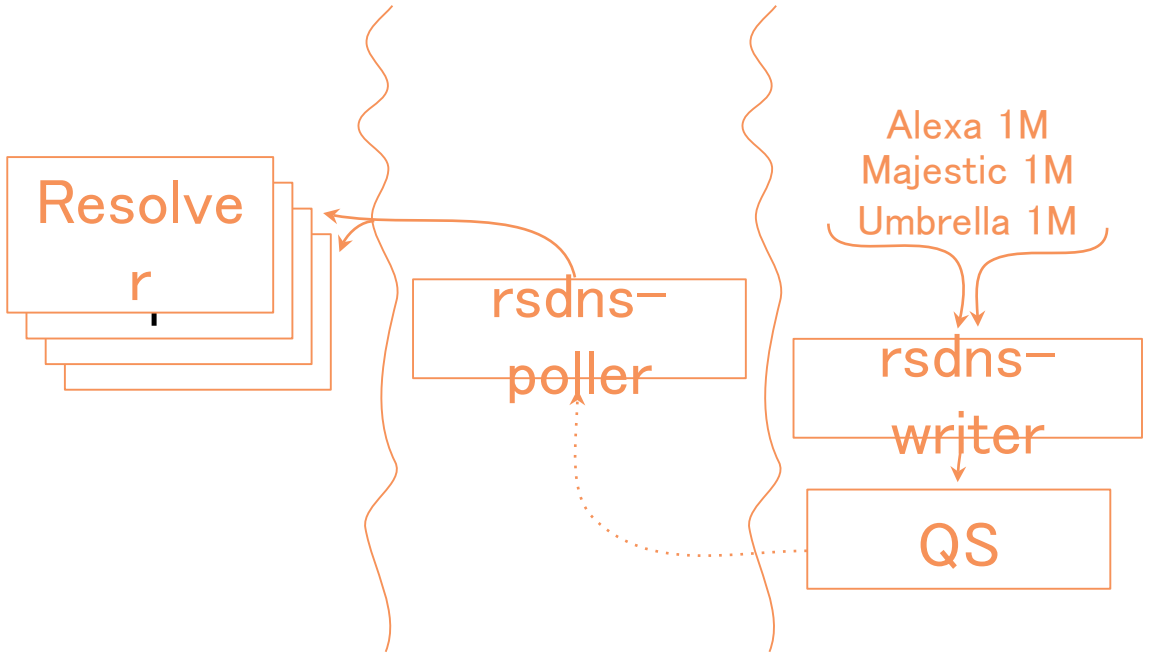
Speed

Speed (prefill)

1.1.1.1

We prefill all caches based on popular domains in a region

- Why: To improve perceived speed and availability
- Popular domains should always be cached
- What is popular?



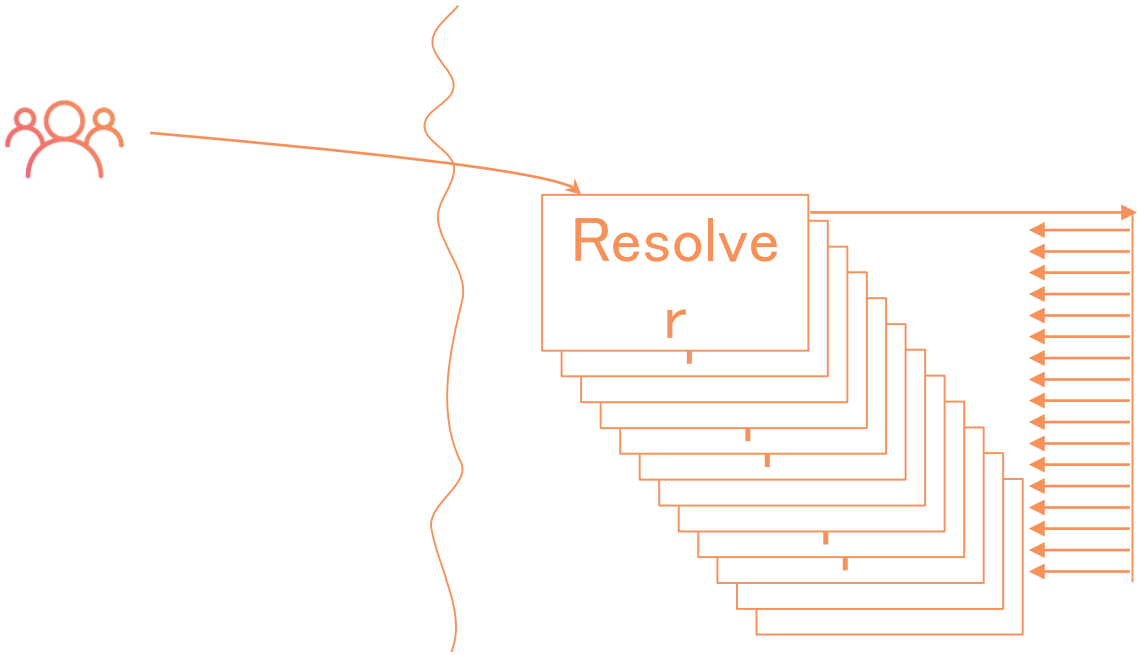
CLOUDFLARE

Speed (backend multicast)

1.1.1.1

Multicasted cache data across machines within the same data center

- Why: Cache hit ratio goes down with the network size
- Cache hit ratio is everything
- Basically a pub-sub
- Consistent latency



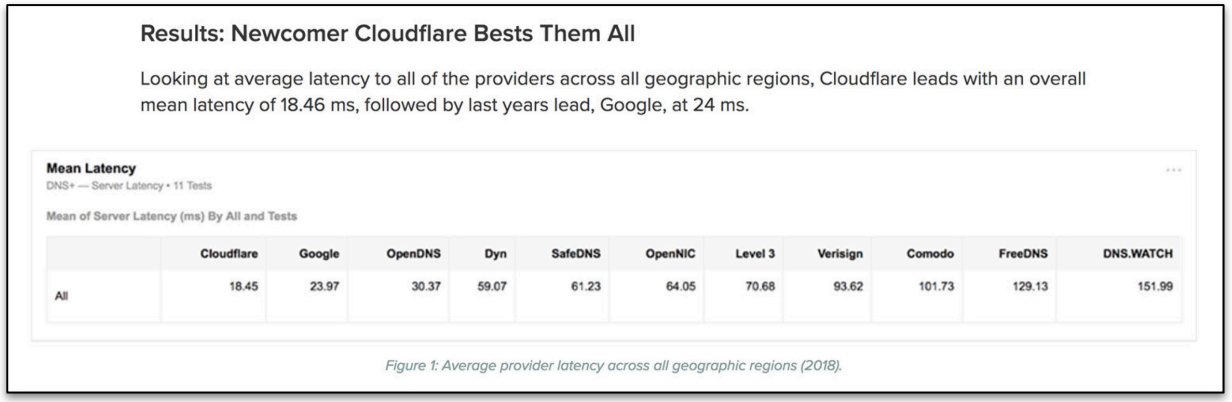
CLOUDFLARE

Speed

1.1.1.1

https://www.dnsperf.com/#!dns-resolvers

	DNS name	Query Speed
1	1.1.1.1	10.24 ms
2	OpenDNS/Umbrella	19.63 ms
3	Quad9	32.45 ms
4	Google	33.97 ms
5	Neustar	45.66 ms
6	Norton	47.46 ms
7	SafeDNS	51.19 ms
8	Verisign	72.24 ms
9	Comodo	82.42 ms
10	Yandex	126.72 ms



Results: Newcomer Cloudflare Bests Them All

Looking at average latency to all of the providers across all geographic regions, Cloudflare leads with an overall mean latency of 18.46 ms, followed by last years lead, Google, at 24 ms.

Mean Latency

DNS+ — Server Latency • 11 Tests

Mean of Server Latency (ms) By All and Tests

	Cloudflare	Google	OpenDNS	Dyn	SafeDNS	OpenNIC	Level 3	Verisign	Comodo	FreeDNS	DNS.WATCH
All	18.45	23.97	30.37	59.07	61.23	64.05	70.68	93.62	101.73	129.13	151.99

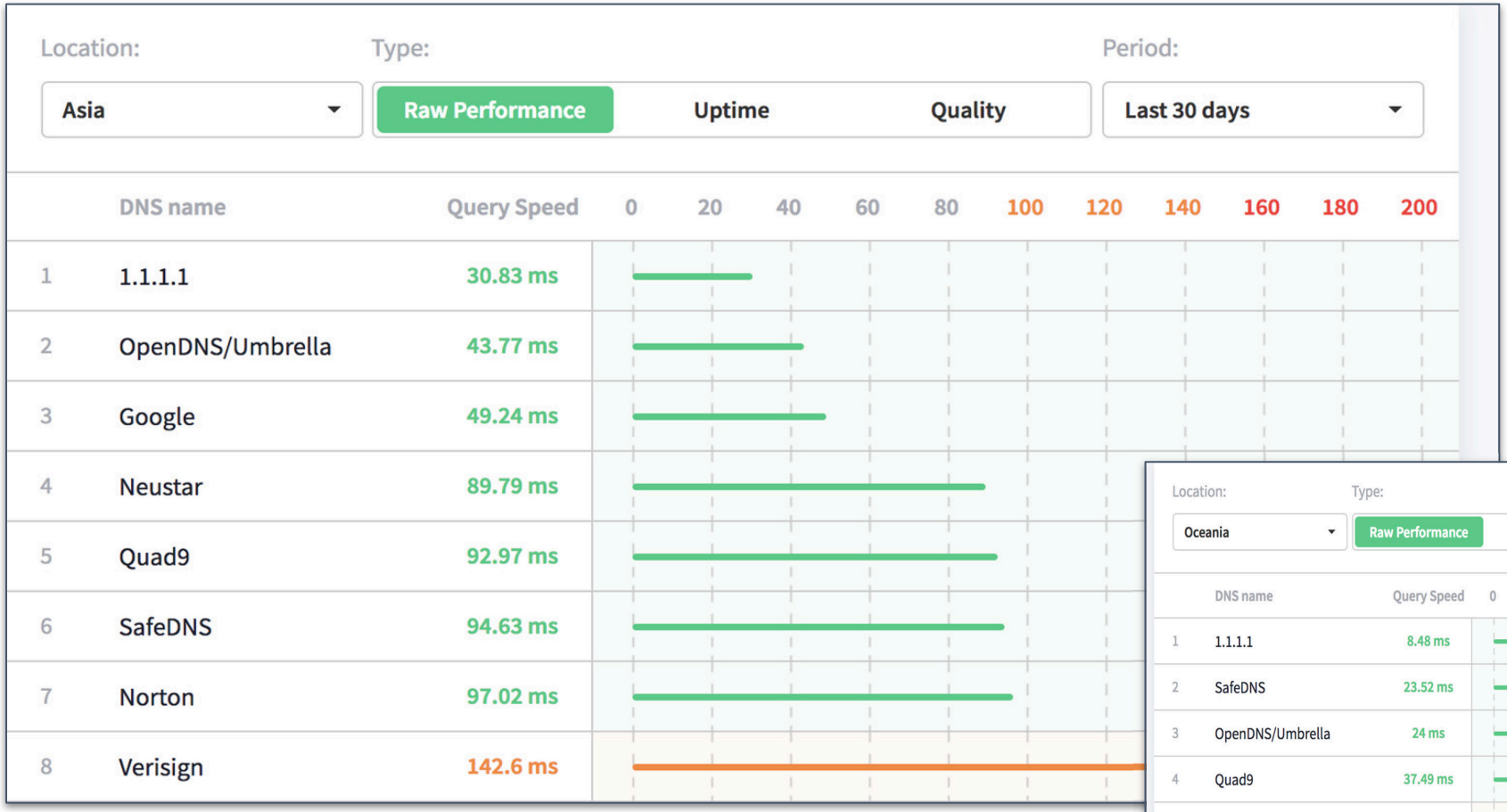
Figure 1: Average provider latency across all geographic regions (2018).

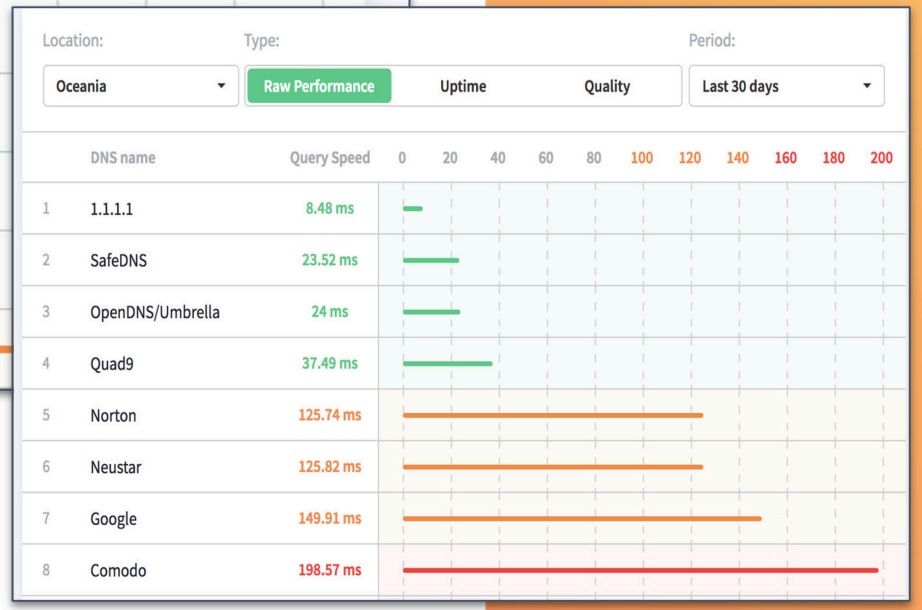
https://blog.thousandeyes.com/ranking-performance-public-dns-providers-2018/

CLOUDFLARE

Speed (in APNIC region)

DNSPerf measurements





CLOUDFLARE

Summary

Summary

- Easy to remember IP addresses
- Support for DOH (DNS over HTTPS) and DNS over TLS
- Cleaning up routing and CPE devices
- Did I mention it’s fast?

1.1.1.1

Setting up the resolver:

https://1.1.1.1/

CLOUDFLARE

Where to from here?

- The blocks are clearly still very tainted
  - Cloudflare continue to sink huge amounts of unrequested, unquenchable traffic
  - Remediation is more than theoretically possible, but its not free
    - Labour costs to help ISPs
    - Those code dependencies embedded in systems in the CPE..
    - We’re talking two /24. Impact on overall IPv4 availability is low
- Public DNS services are now quite popular
  - Google pDNS
  - IBM/PCH 9.9.9.9 “cleanfeed”
  - CloudFlare 1.1.1.1 has high adoption rate
- Continuance with renewal unless clear policy drivers dictate otherwise
  - APNIC Labs does not propose a change at this time

Questions?