Client-Side IPv6 Measurement at Scale

Geoff Huston
APNIC Labs

How to measure millions of end devices for their IPv6 capability?

How to measure millions of end devices for their IPv6 capability?

Be

How to measure millions of end devices for their IPv6 capability?

OR

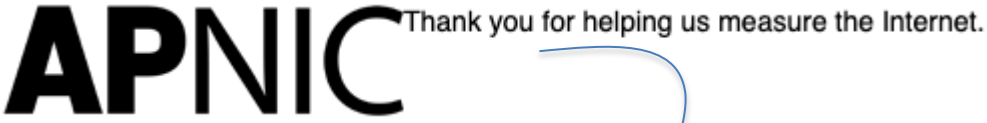
Have your measurement code run on millions of end devices

APNIC's Approach

- we wanted to measure IPv6 deployment as seen by end users
- We wanted to say something about ALL users
- So we were looking at a way to sample end users in a random but statistically significant fashion
- We stumbled across the advertising networks...

APNIC

Thank you for helping us measure the Internet.




```
<!DOCTYPE html>
<html><head data-gwd-animation-mode="quickMode"><meta name="GCD"
content="YTk3ODQ3ZWZhN2I4NzZmMzBkNTEwYjJl657daa7a9fa4c339ce298ace1f626e3e"/>

  <meta name="generator" content="Google Web Designer 1.2.1.0121">
  <meta http-equiv="Content-Type" content="text/html; charset=utf-8">
  <meta name="viewport" content="width=device-width, initial-scale=1.0">

  <script type="text/javascript" src="https://s0.2mdn.net/ads/studio/Enabler.js"></script>
  </script>
  <script type="text/javascript" src="html5ad.js">
 </script>
<body>
  <img src="apnic-logo.png" style="float:left"/>
  Thank you for helping us measure the Internet.

  <script type="text/javascript">
    runLabsTests();
  </script>
  <!-- This section contains metadata about the ad. Most importantly, the ad size. -->
  <script type="text/gwd-admetadata">
{"version":1,"type":"DoubleClick","format":"","template":"","politeload":true,"counters":[],"timers":[],"exits":
[],"creativeProperties":{"minWidth":468,"minHeight":60,"maxWidth":468,"maxHeight":60},"components":[]}</script>
</body></html>
```

The Ad Measurement Technique

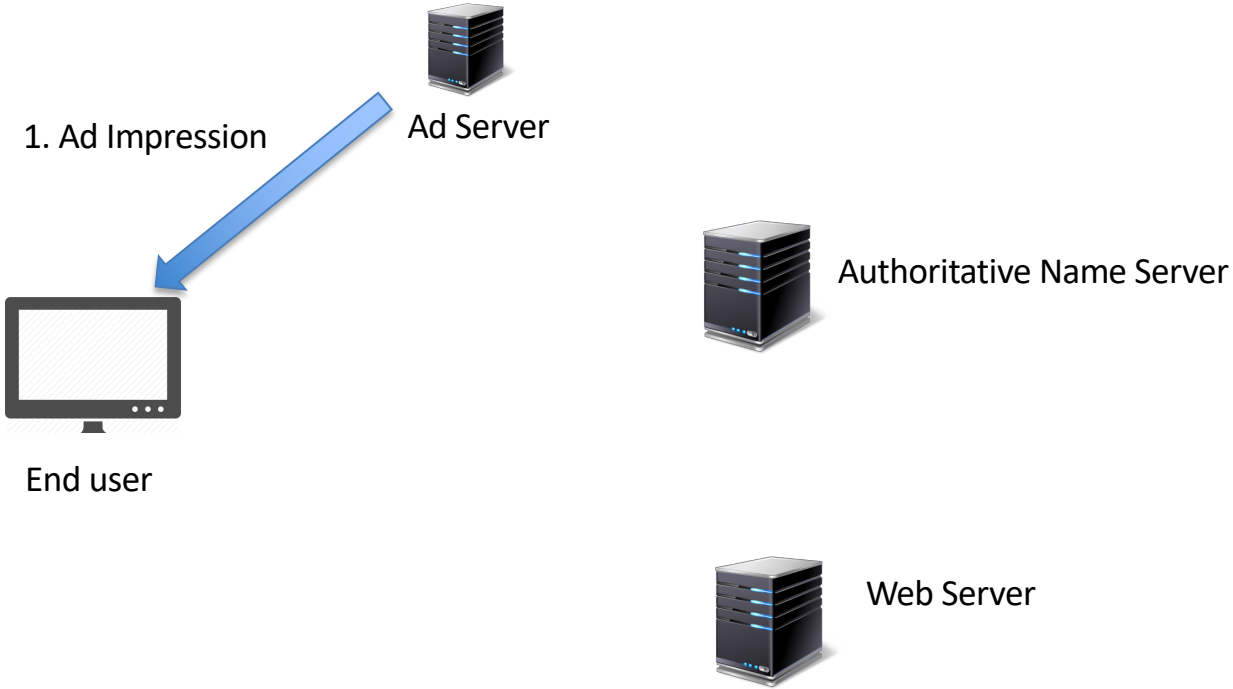
Ad Server

Authoritative Name Server

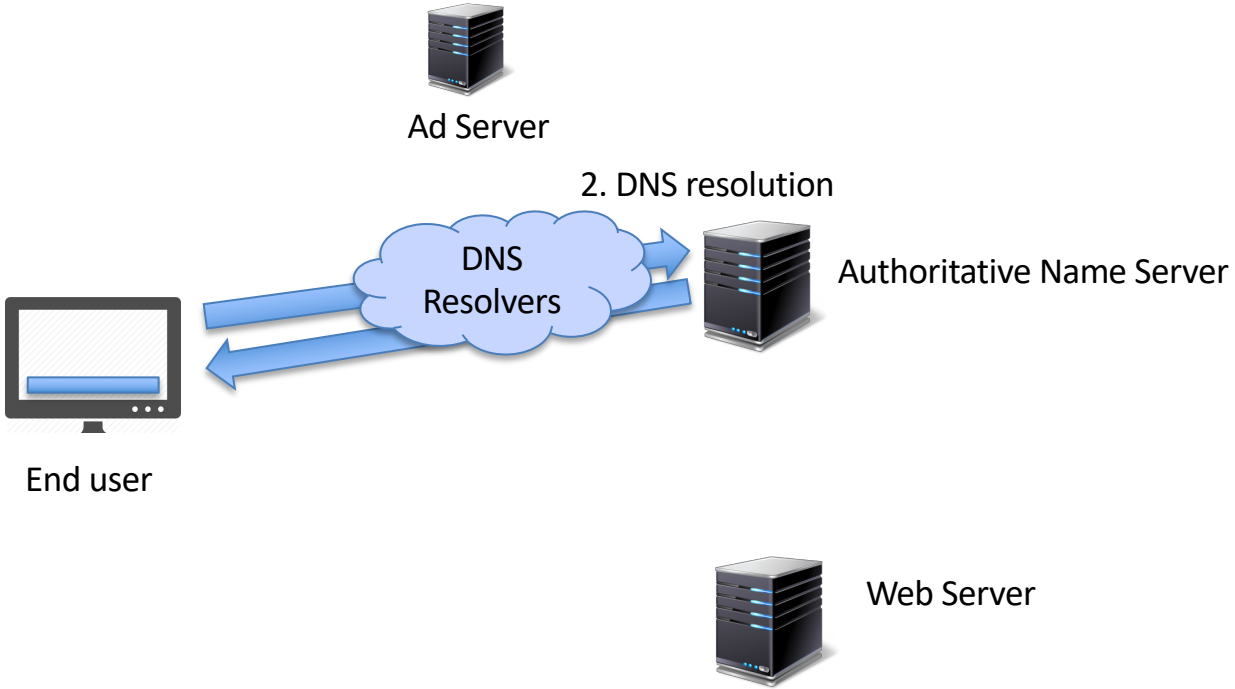
End user

Web Server

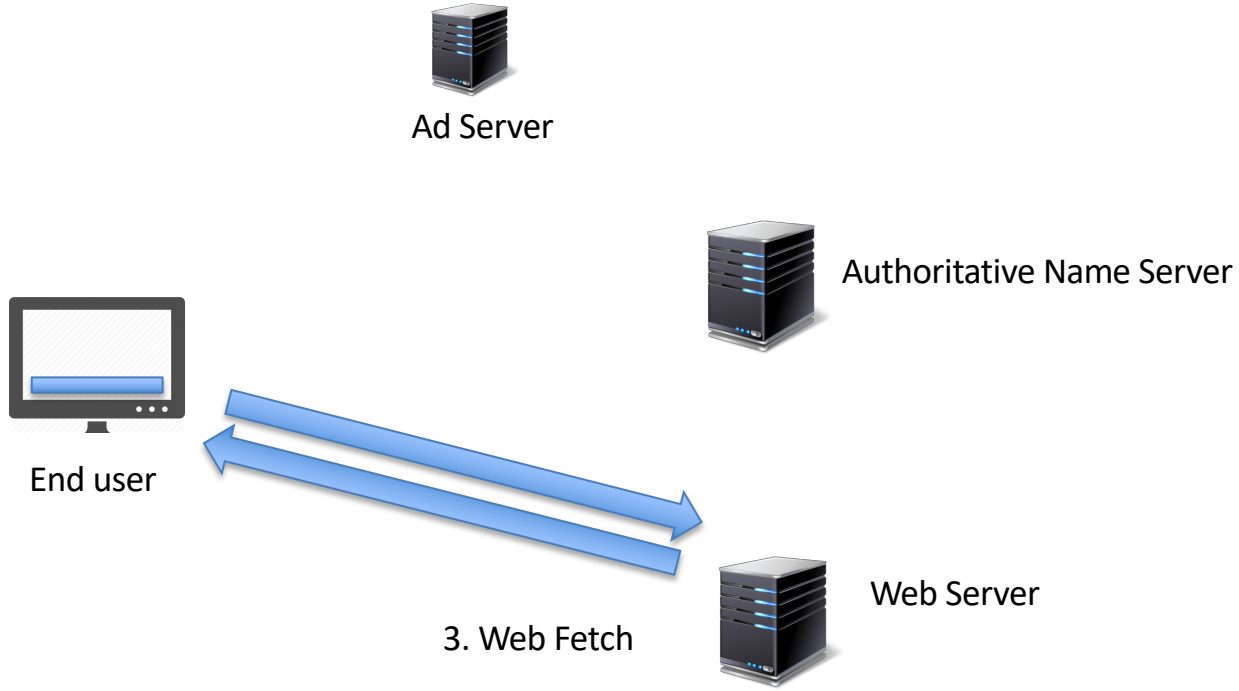
The Ad Measurement Technique



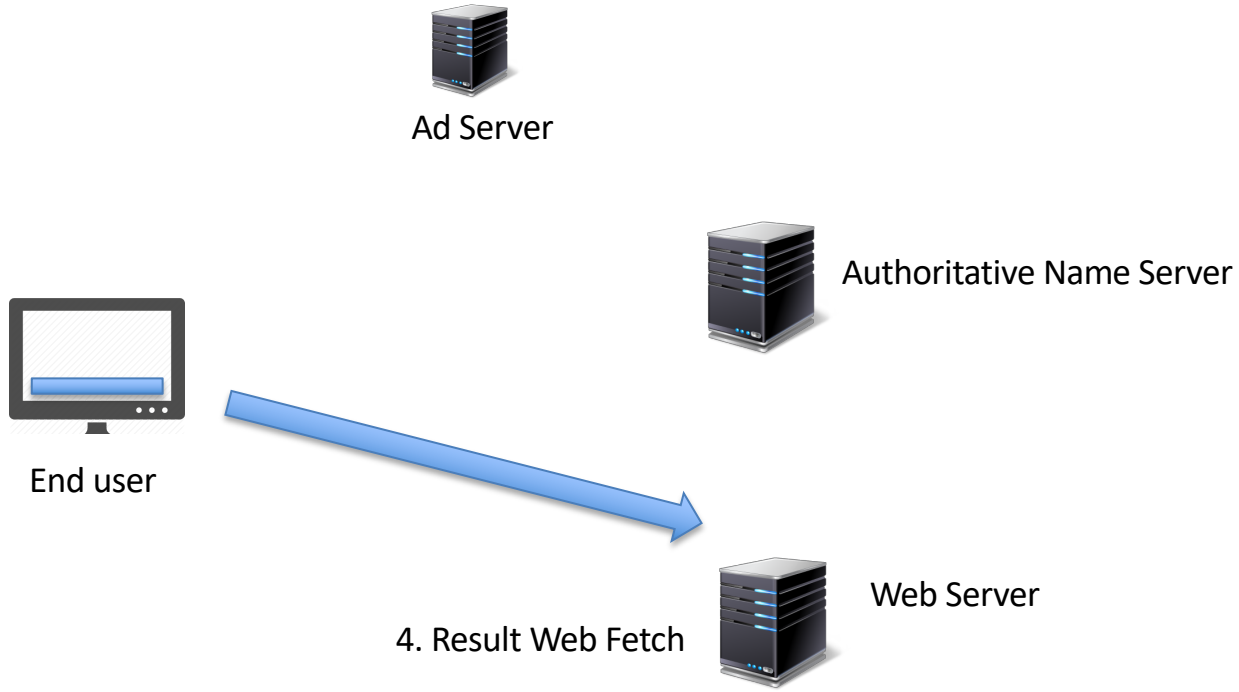
The Ad Measurement Technique



The Ad Measurement Technique



The Ad Measurement Technique



What can be scripted

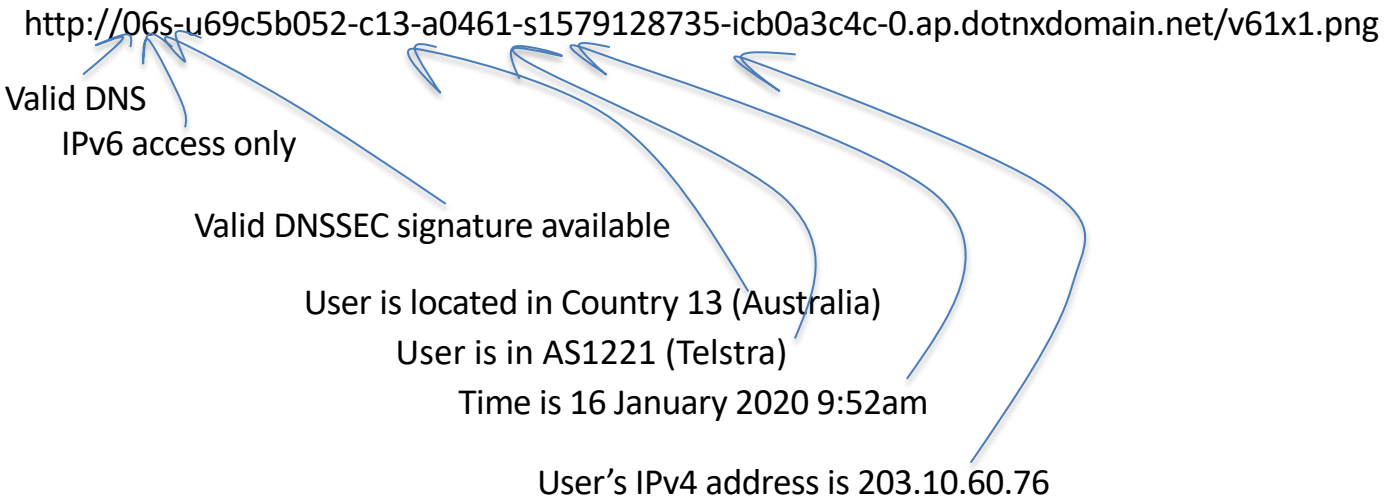
- Not much:
  - http.FetchImg()

    i.e. attempt to retrieve a URL

- But that's enough!
  - It's EXACTLY what users do!
  - A URL consists of a DNS question and an HTML question
  - What if we point both the DNS and the HTML to servers we run?
  - As long as each Ad execution uses unique names we can push the user query back to our servers

Tests

Think of a URL name as a microcoded instruction set directed to programmable DNS and HTTP servers …



Tests

- We listen on /64s – so we can use nonce IPv6 addresses
- Currently we are looking at encoding IPv4 and time in the IIF to detect replay activities

Ad Placement

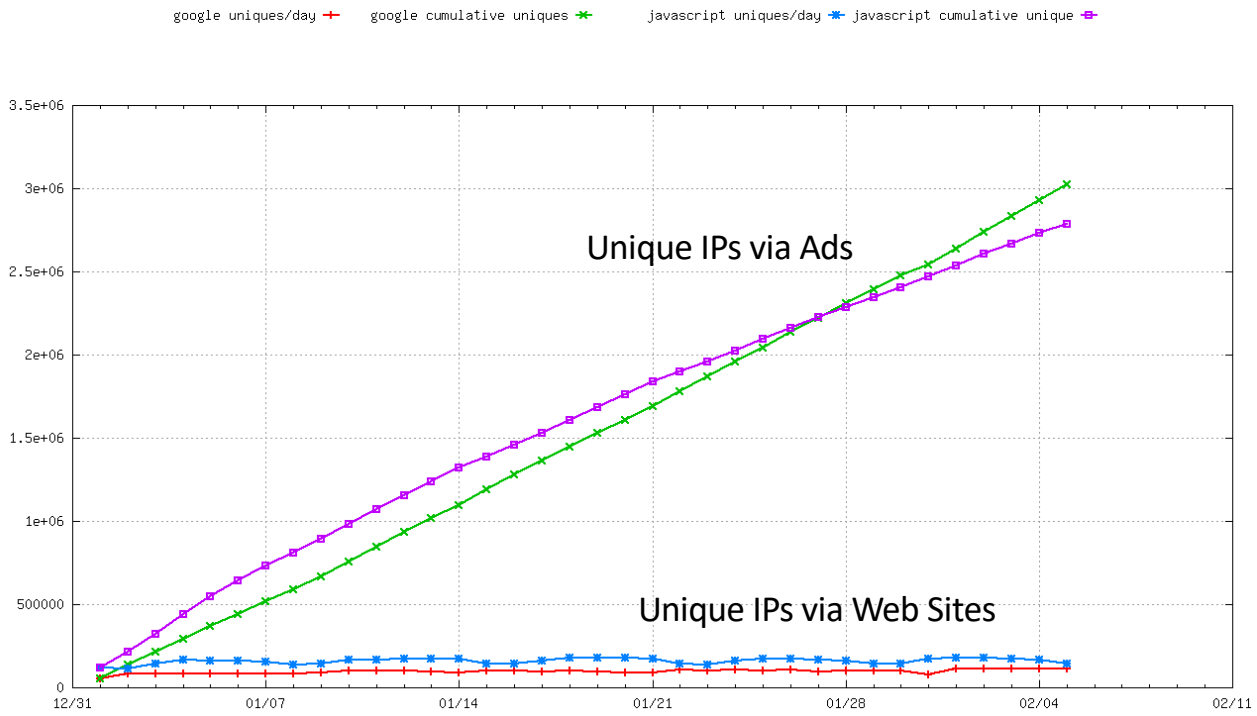
At low CPM, the advertising network needs to present unique, new eyeballs to harvest impressions and take your money.

- Therefore, a ‘good’ advertising network provides fresh crop of unique clients per day

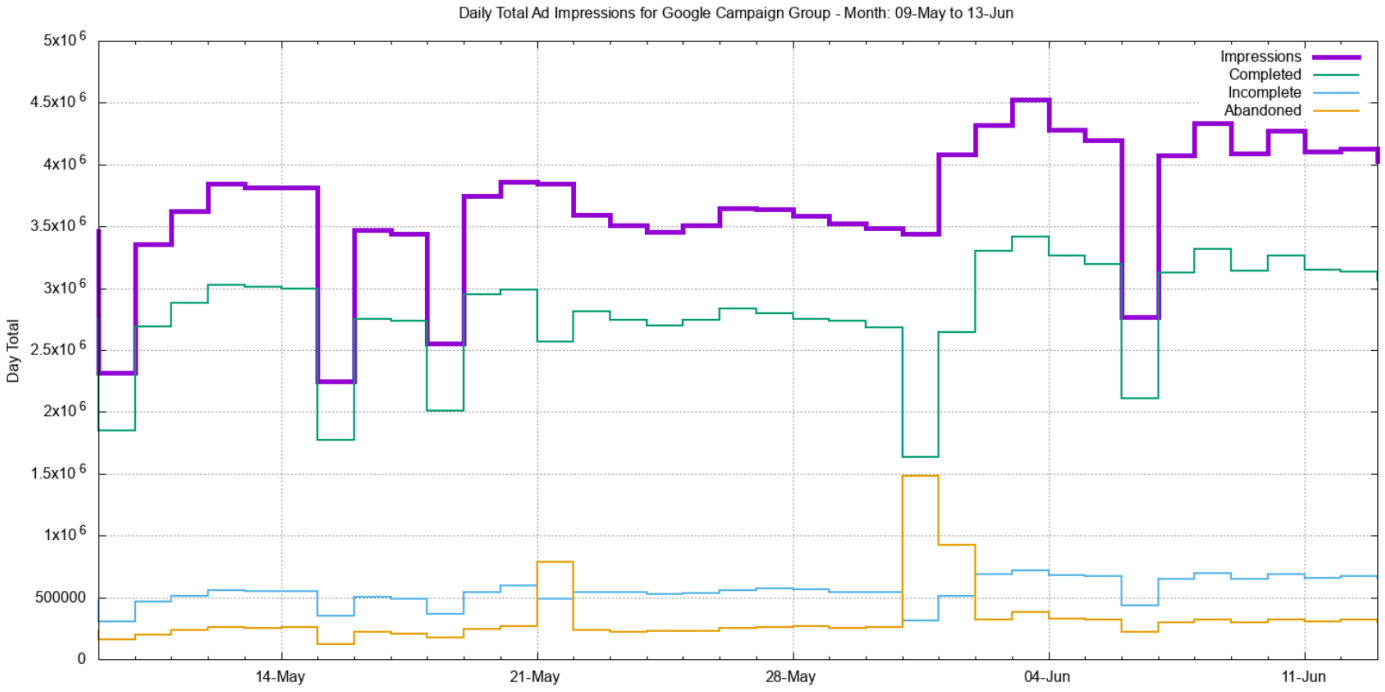
Unique IPS?

- Collect list of unique IP addresses seen
  - Per day
  - Since inception
- Plot to see behaviours of system
  - Do we see ‘same eyeballs’ all the time?

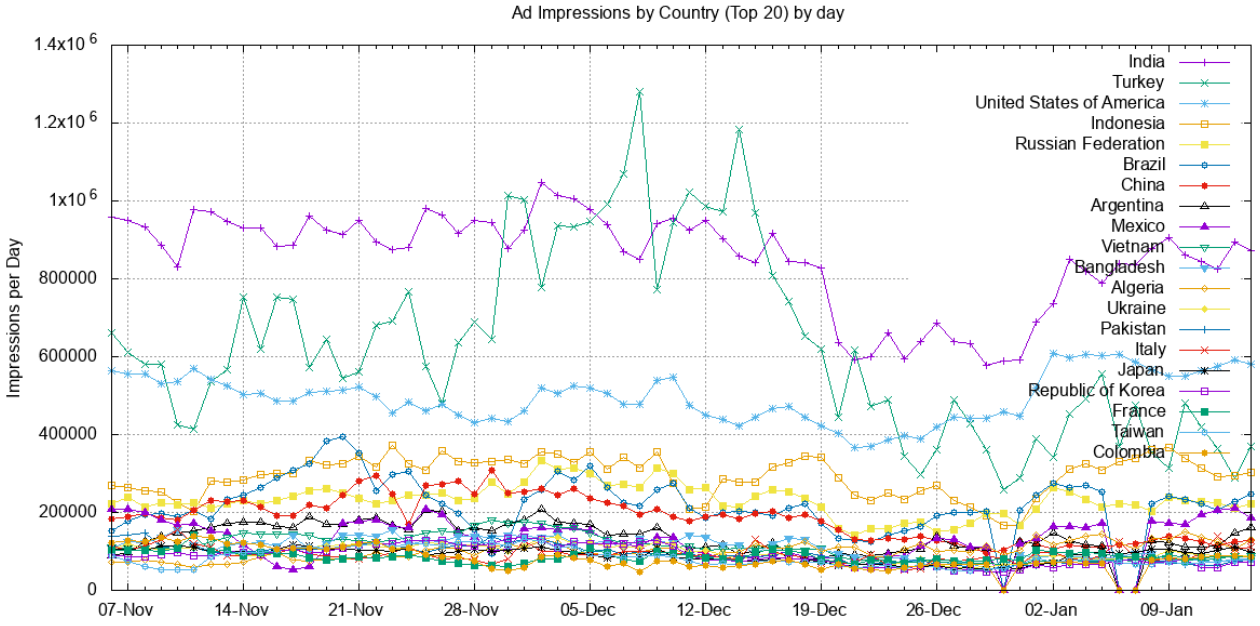
Lots of Unique IP'S



Ad Presentation Volumes

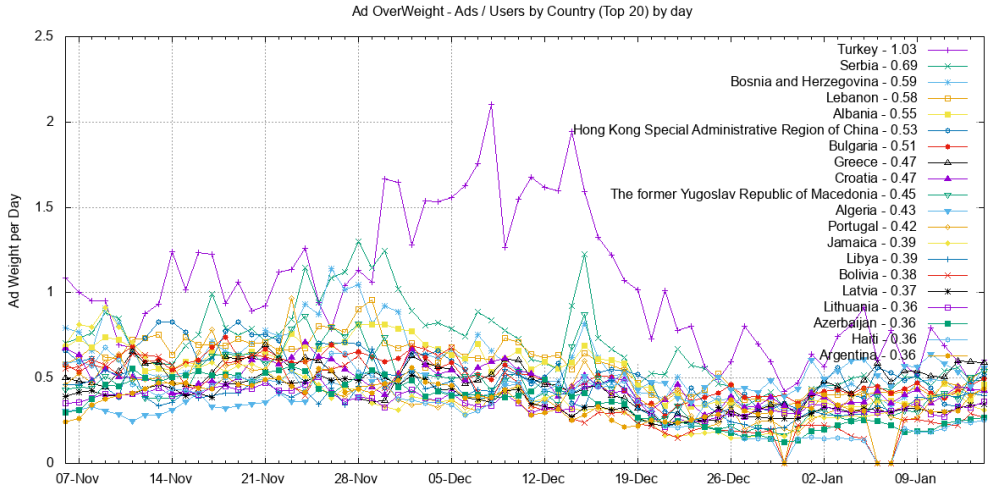


Ad Presentations: Countries



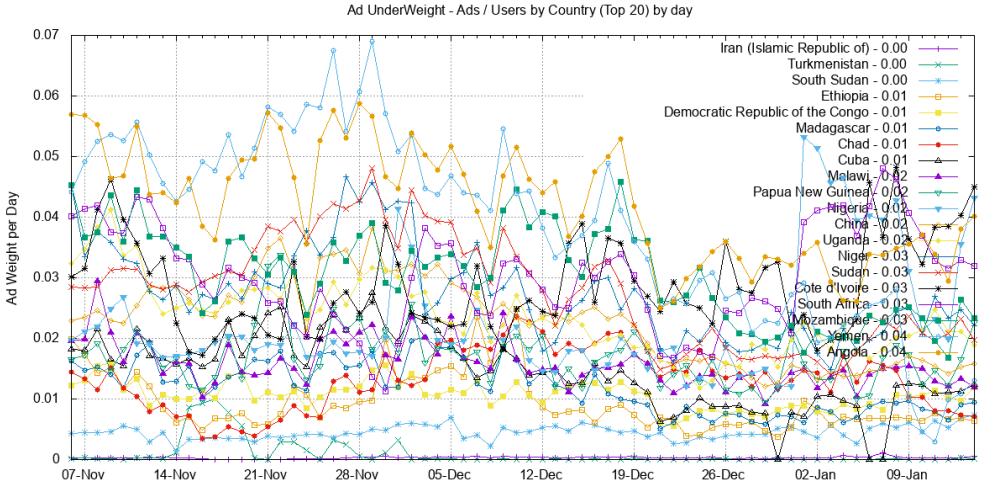
Bias Compensation

- The ad presentation is NOT uniform across the Internet's user population
  - The ad machinery 'over-presents' in some countries:



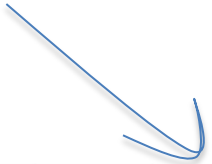
Bias Compensation

- The ad presentation is NOT uniform across the Internet's user population
  - The ad machinery 'under-presents' in some countries:



Bias Compensation

- Use ITU data on Internet users per country as the reference set, and weight the ad results to compensate for ad placement bias



CC	Country	IPv6 Capable	IPv6 Preferred	Samples	Weight	Weighted Samples
IN	India, Southern Asia, Asia	63.64%	62.68%	1,030,515	0.97	1,003,783
BE	Belgium, Western Europe, Europe	58.89%	58.59%	13,139	1.64	21,592
US	United States of America, Northern America, Americas	55.76%	55.06%	782,992	0.78	614,124
MY	Malaysia, South-Eastern Asia, Asia	47.35%	46.80%	55,056	0.85	46,901
DE	Germany, Western Europe, Europe	46.12%	45.43%	78,484	1.95	153,196
GR	Greece, Southern Europe, Europe	45.69%	45.49%	56,917	0.27	15,228
TW	Taiwan, Eastern Asia, Asia	44.58%	43.06%	95,145	0.42	39,994

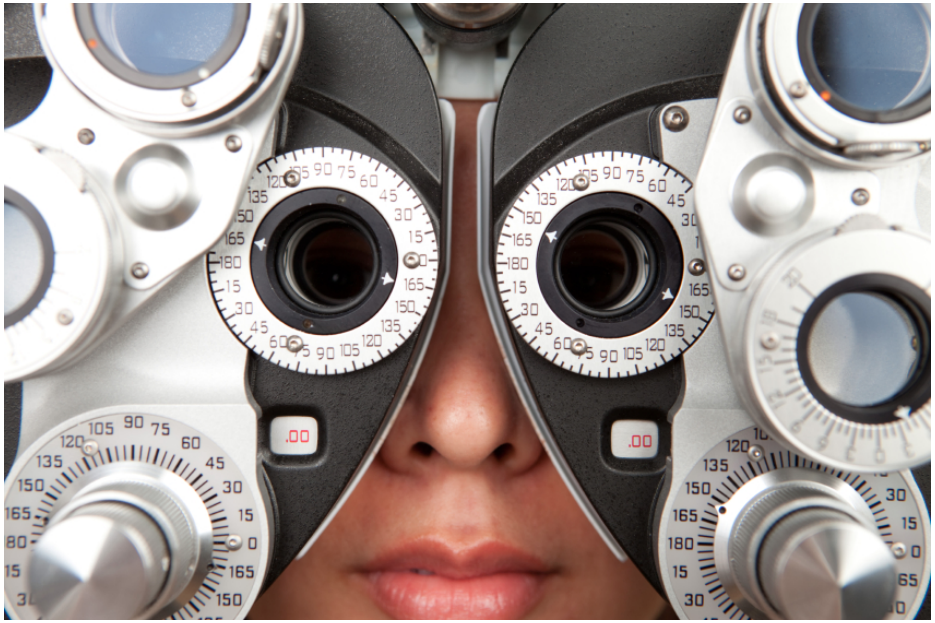
Dealing with the data

- Unified web logs, dns query logs, packet capture
- Map individual DNS and HTML transactions using a common experiment identifier
- For example:
  - DNSSEC validation implies:
    - DNS queries include EDNS(0) DNSSEC OK flag set
    - See DNS queries for DNSSEC signature records (DNSKEY / DS)
    - User fetches URL corresponding to a validly signed DNS name
    - User does not fetch URL corresponding to a in validly signed DNS name

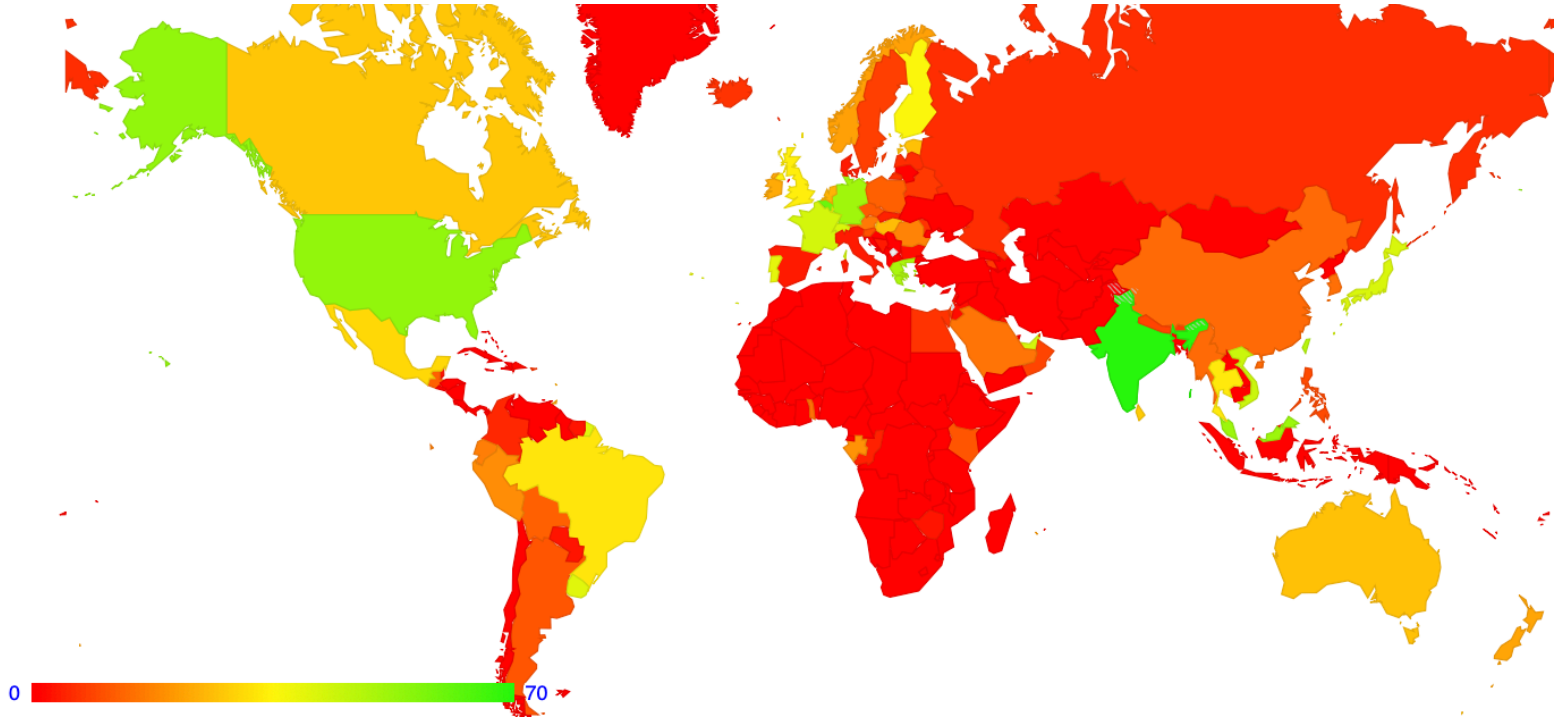
What are we measuring?

- IPv6 Adoption
- IPv6 Dual Stack Preference
- IPv6 Performance
- IPv6 FragmentationExtension header fragility
-

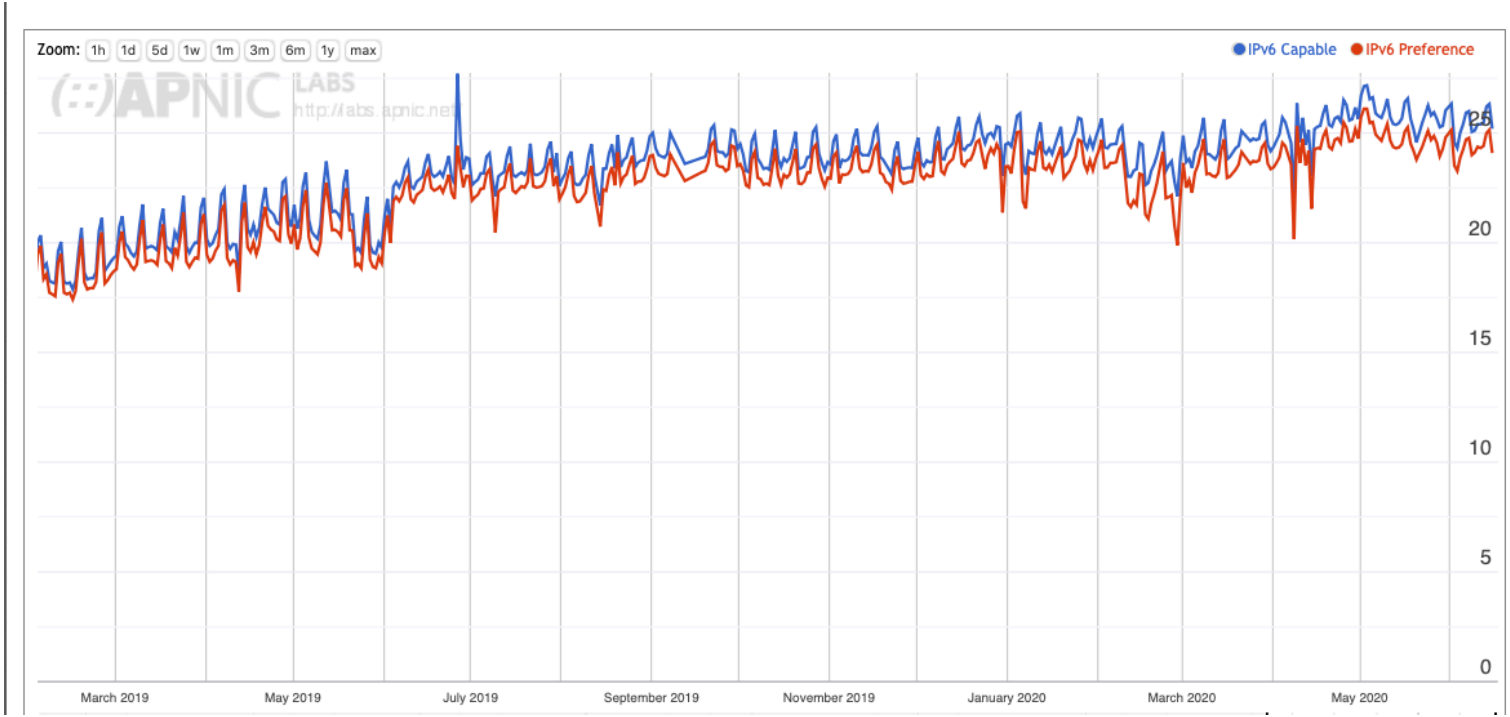
What are we seeing?



IPv6 Adoption by Country

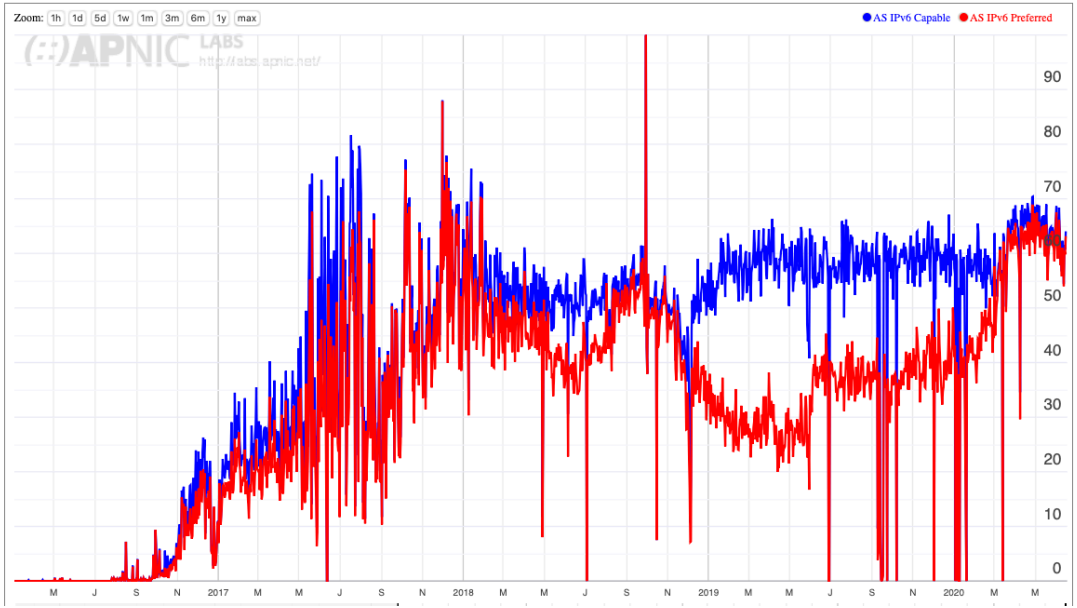


IPv6 Adoption and Preference

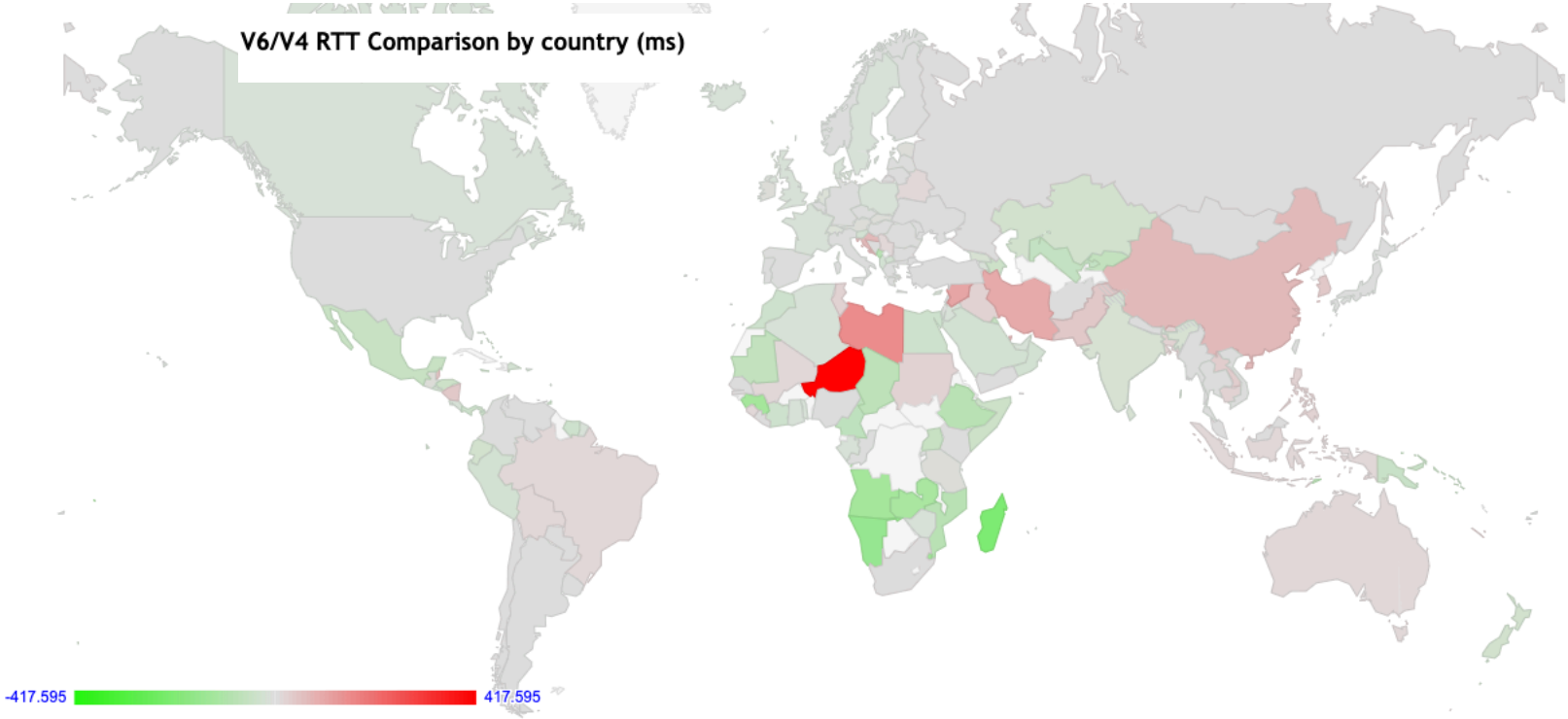


IPv6 Preference

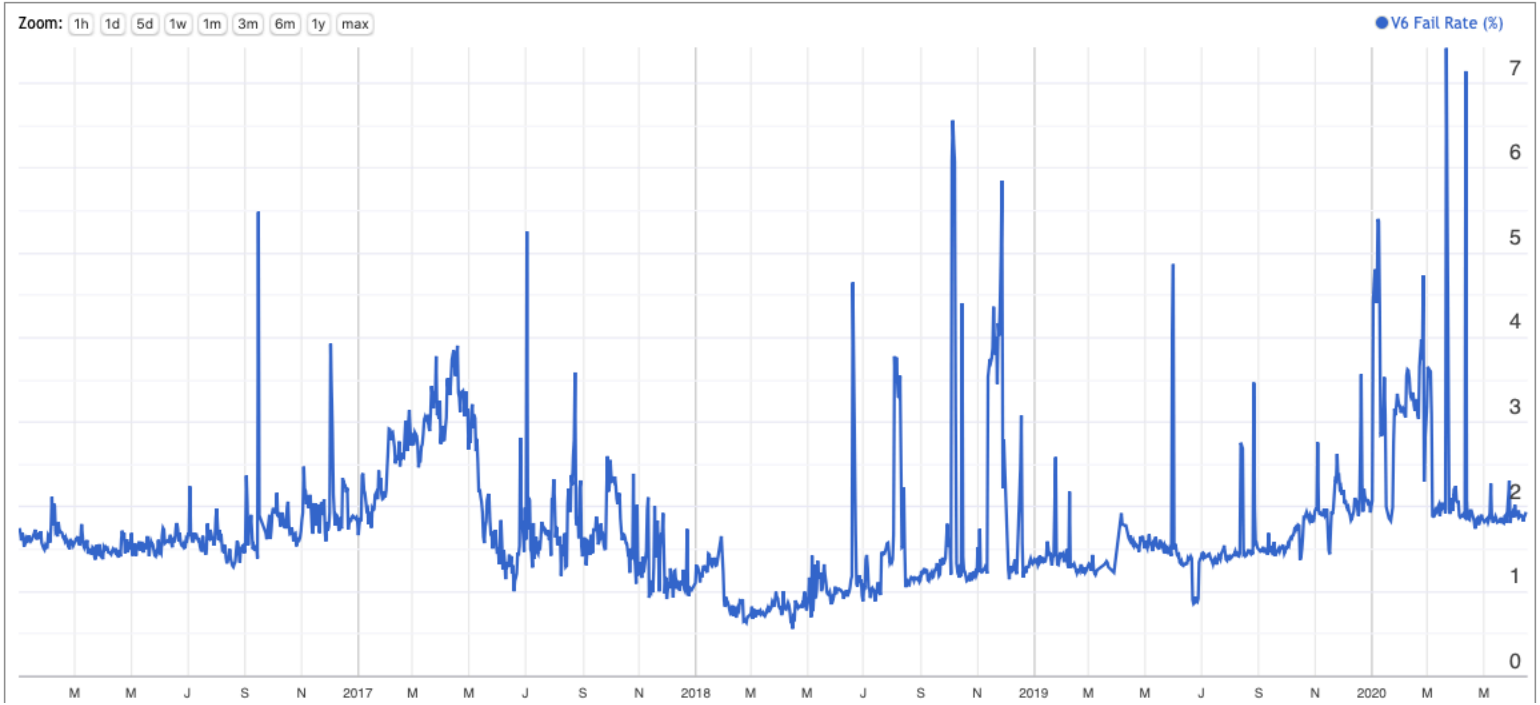
IPv6 Per-Country Deployment for AS9500: VODAFONE-TRANSIT-AS Vodafone NZ Ltd., New Zealand (NZ)



IPv6 Performance



IPv6 Reliability



But…

It’s not a general purpose compute platform, so it can’t do many things

- Ping, traceroute, etc
- Send data to any destination
- Pull data from any destination
- Use different protocols

This is a “many-to-one” styled setup where the server instrumentation provides insight on the inferred behaviour of the edges

Measurement Ethics

- There is no user consent
- And cookies (even “don’t measurement me!” cookies) are progressively being frowned upon by the ad placement folk

Which means we:

- Don’t generate large data volumes
- Don’t use ‘compromising’ URL names
- Don’t publish PII
- Don’t share any raw data from our DNS and HTTP servers

In Summary…

- Measuring what happens at the user level by measuring some artifact or behaviour in the infrastructure and inferring some form of user behaviour is always going to be a guess of some form
- If you really want to measure user behaviour then its useful to trigger the user to behave in the way you want to study or measure
- The technique of embedding simple test code behind ads is one way of achieving this objective
  - for certain kinds of behaviours relating to the DNS and to URL fetching