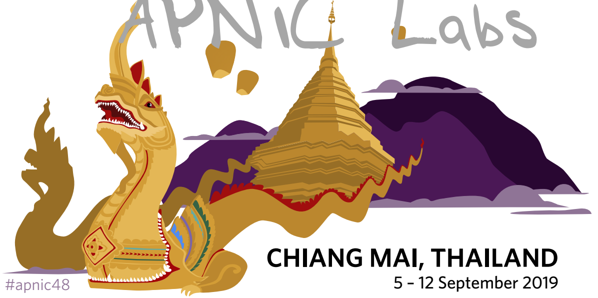
APNIC 48

IPv6 Reliability Measurements

Geoff Huston
APNIC Labs



#apnic48

CHIANG MAI, THAILAND
5 - 12 September 2019

The Question is…

Does IPv6 perform “better” than IPv4?

The Question is…

Does IPv6 perform "better" than IPv4?

Is it more reliable?

Is it faster?

The Measurement

- The measured endpoint retrieves two URLs from the same remote server – one using IPv4 and the other using IPv6
  - Unique DNS names and TLS are used to ensure that caching does not play a role in the measurement
- We perform full packet capture at the server
- We measure some 12M sample points each day and some 22% of these measurements use both IPv4 and IPv6

Analysis

- We look at the SYN/ACK exchange at the start of the TLS session
- The time between receipt of the SYN and the subsequent ACK at the server is no less than one RTT between the server and the endpoint (and is a reasonable first order substitute for an RTT)
- A received SYN with no subsequent ACK is interpreted as a failed connection attempt

Analysis - Reliability

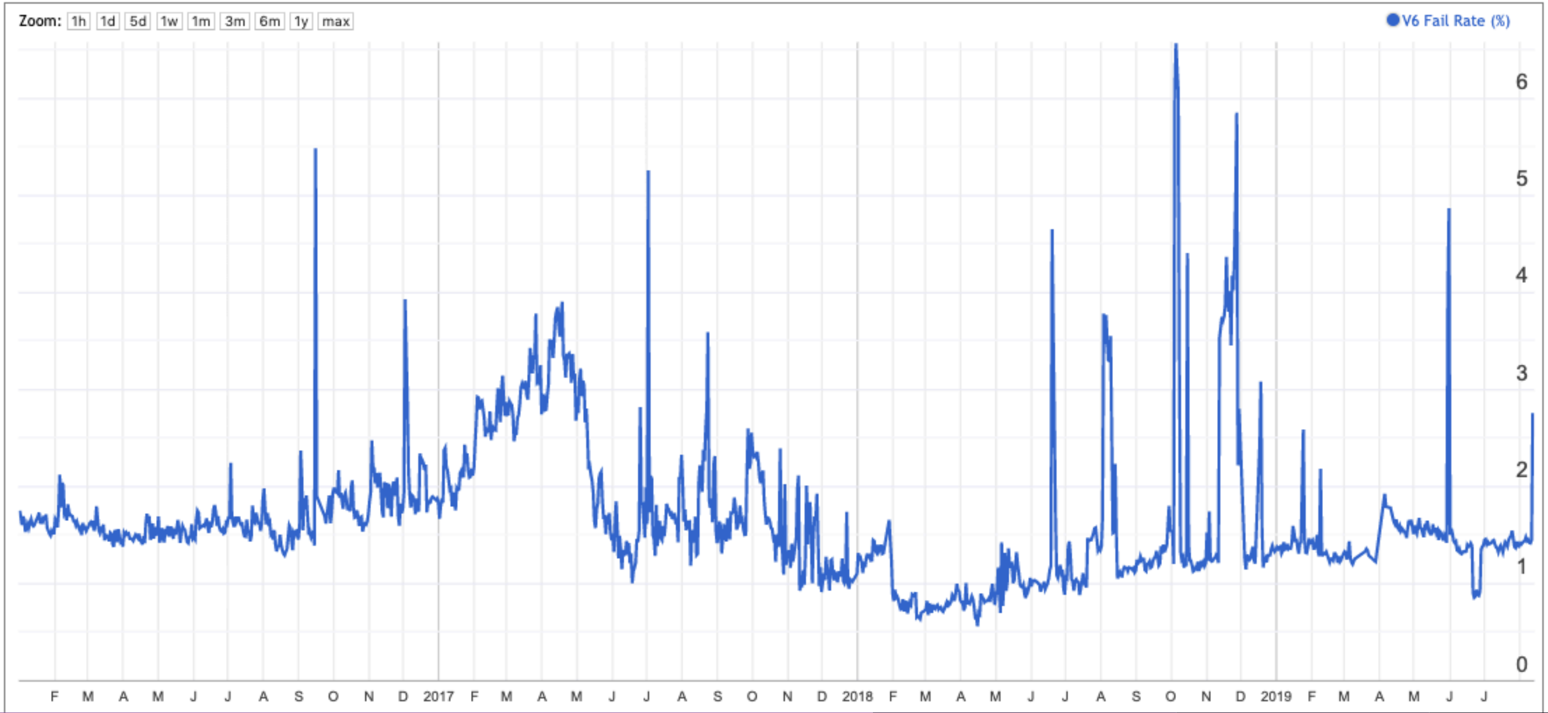
- Why measure SYN handshake failure?
- In a dual stack environment many of the most widely used apps (browsers) use Happy Eyeballs to decide which protocol to select
- Happy Eyeballs bases its decision on the first protocol to complete a TCP SYN handshake
- So TCP handshake failure will strongly influence this decision

Analysis - Speed

- Why measure only the handshake delay? Why not measure a larger data transfer?
- Because in the end host and the server the same TCP version is used on top of IPv4 and IPv6
  – If the end to end paths are the same in IPv4 and IPv6 we would see precisely the same session throughput
- RTT and packet loss probability determine session throughput
  – In this experiment we use the RTT as in indicator of path difference

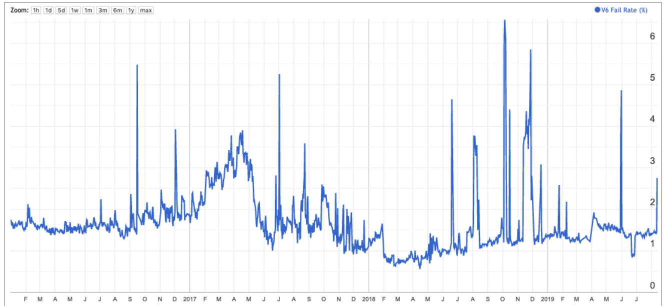
IPv6 TCP Connection Failure

Average V6 Connection Failure Rate for World (XA)



IPv6 TCP Connection Failure

Average V6 Connection Failure Rate for World (XA)



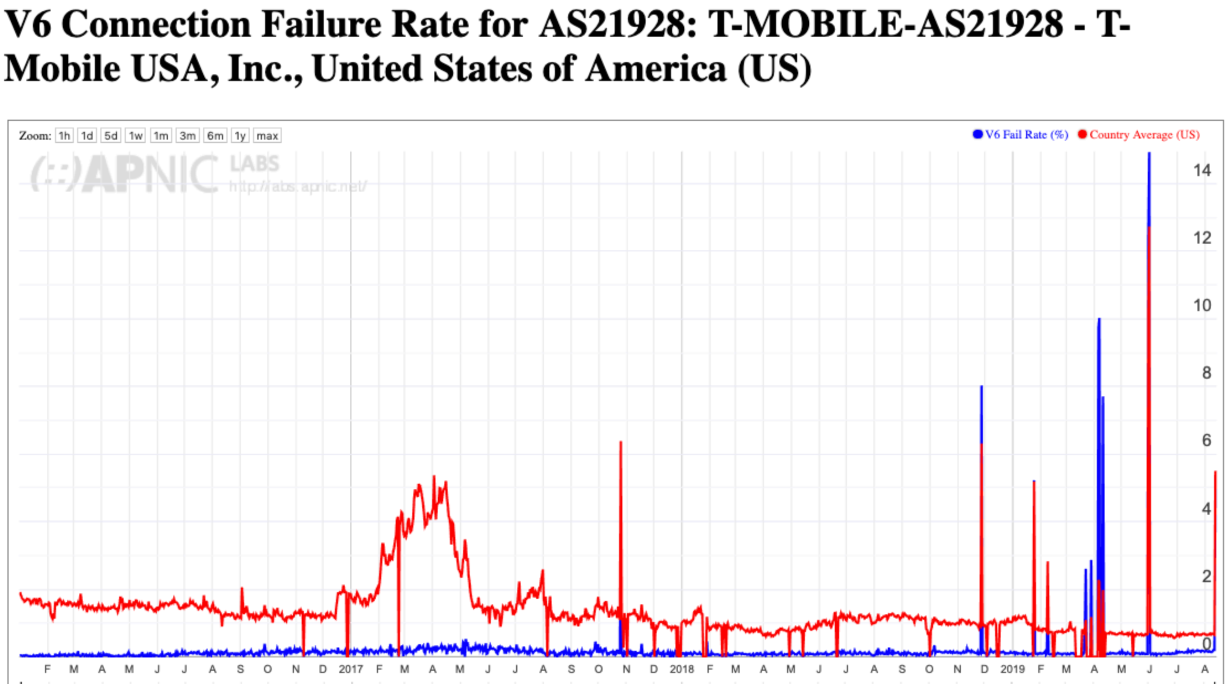
The global failure rate of some 1.4% is better than earlier data (4% failure in early 2017), but its still bad

We cannot detect failure in attempting to deliver a packet from the client to the server – what we see as "failure" is a failure to deliver an IPv6 packet from the server to the client

Possible reasons:

- Endpoint using an unreachable IPv6 address
- End site firewalls and filters
- Transition mechanism failure

The Good



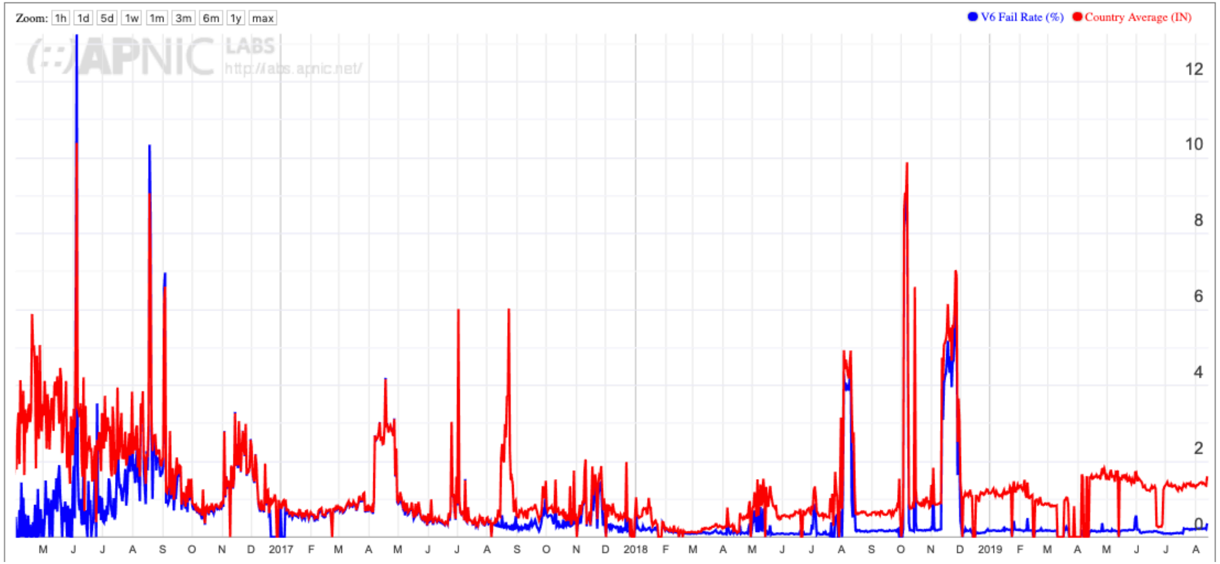
This 464XLAT mobile network (T-Mobile) has remarkably small failure rates – the endpoints are connected via native IPv6 and as this is a mobile network there is only a small amount of customer-operated filtering middleware

464XLAT Performance

- These networks operate in a “native” IPv6 mode
- IPv6 connections to a server require no network processing and no client handling

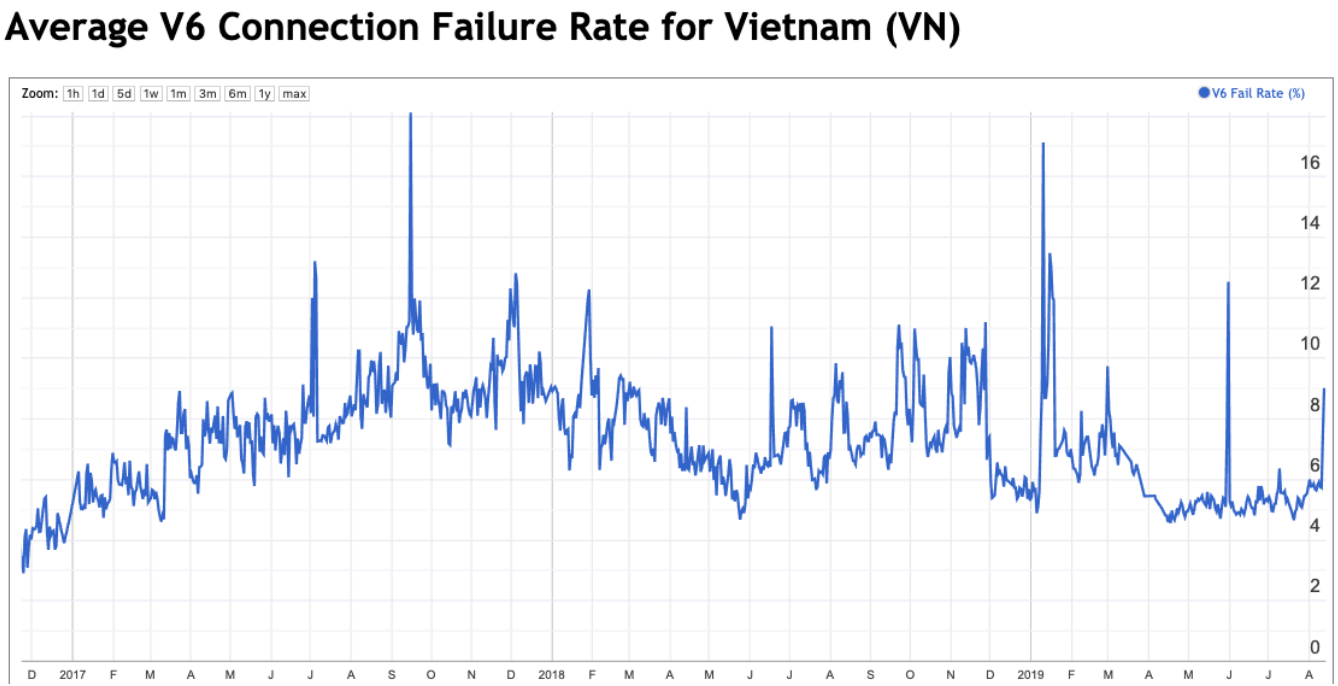
The Good

V6 Connection Failure Rate for AS55836: RELIANCEJIO-IN Reliance Jio Infocomm Limited, India (IN)



Similar story in India with Reliance JIO – the endpoints are connected via native IPv6 and as this is a mobile network there is only a small amount of customer-operated filtering middleware

The Bad

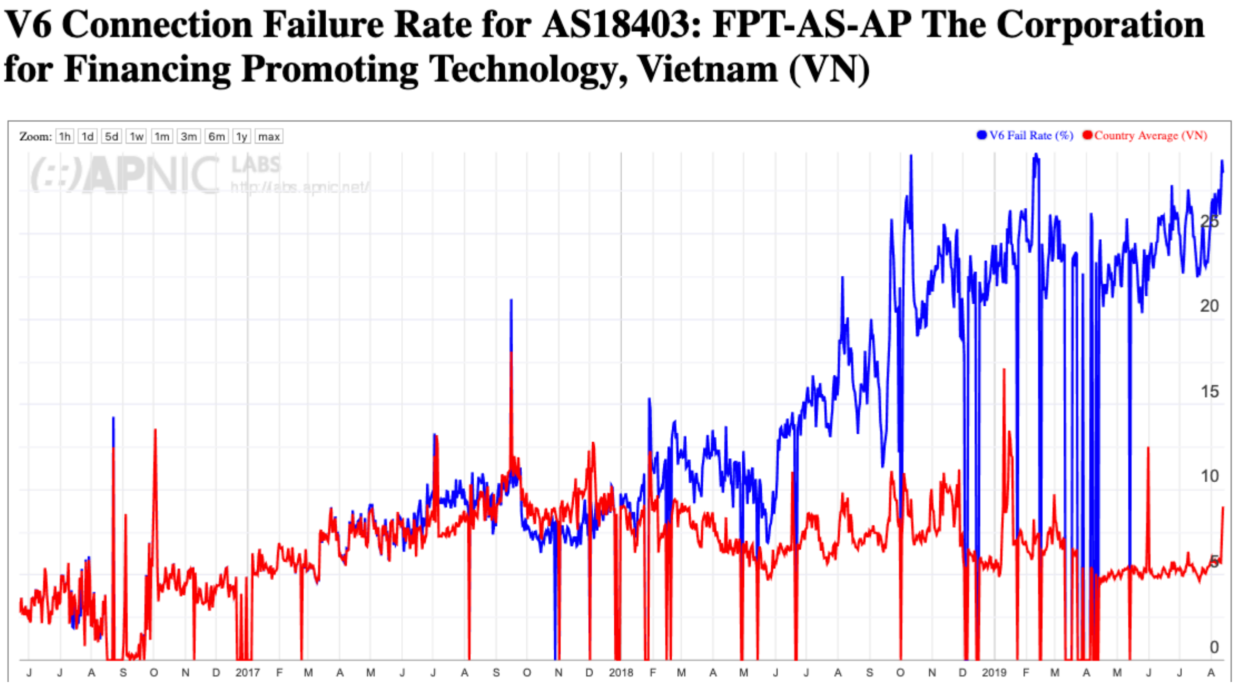


Seriously?

A 6%-10% IPv6 connection failure rate is bad enough

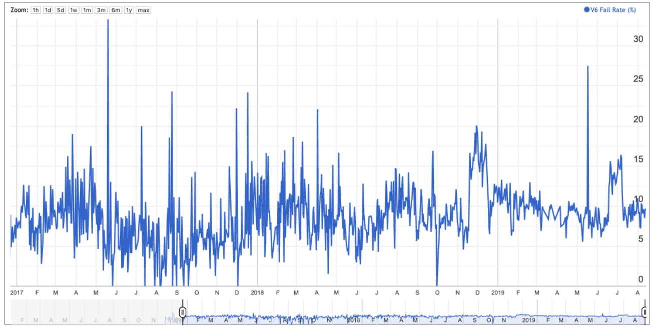
A sustained high failure rate for over 2 years seems worse!

The Appalling!



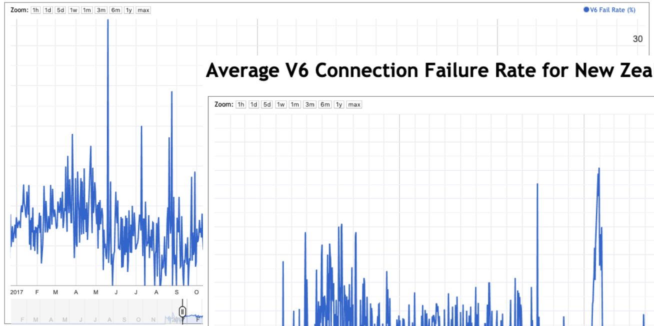
High IPv6 Failure Levels

Average V6 Connection Failure Rate for Guatemala (GT)

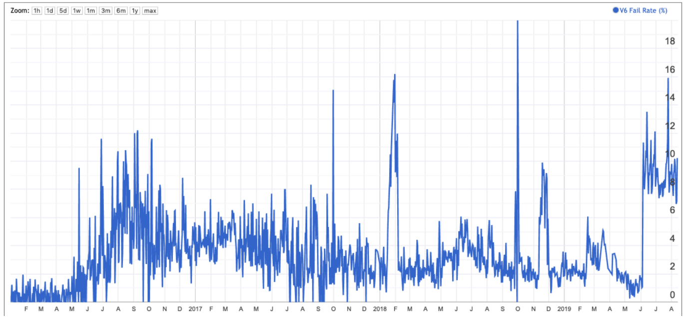


High IPv6 Failure Levels

Average V6 Connection Failure Rate for Guatemala (GT)

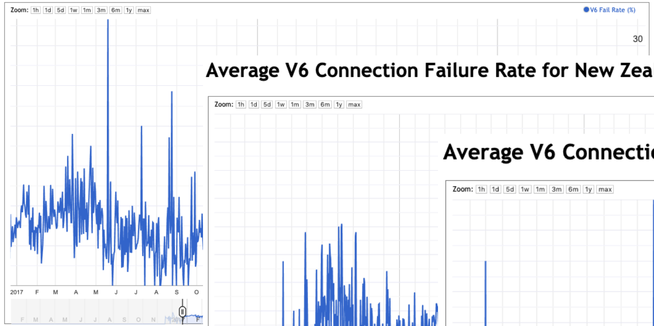


Average V6 Connection Failure Rate for New Zealand (NZ)

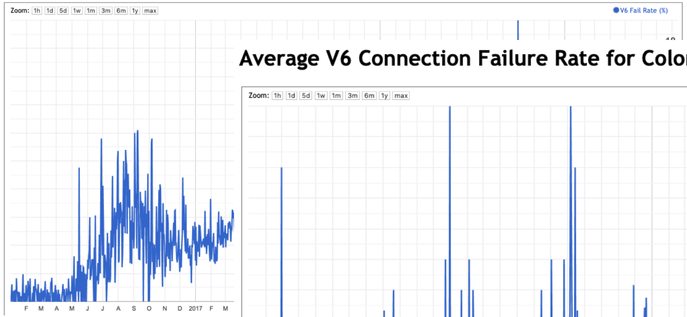


High IPv6 Failure Levels

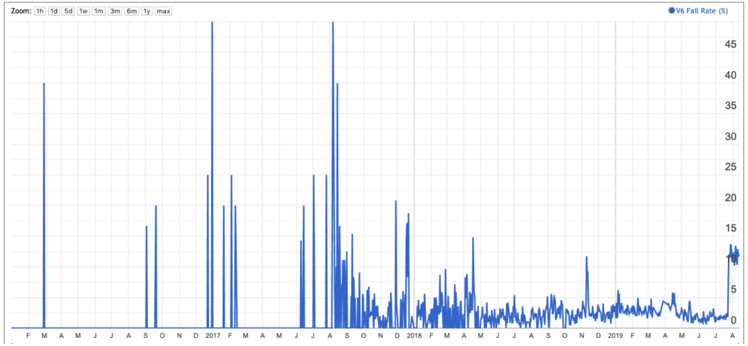
Average V6 Connection Failure Rate for Guatemala (GT)



Average V6 Connection Failure Rate for New Zealand (NZ)

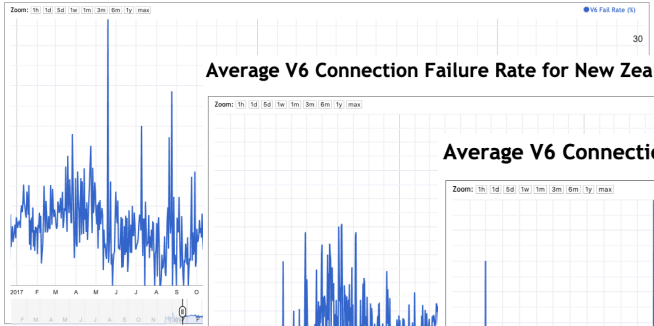


Average V6 Connection Failure Rate for Colombia (CO)

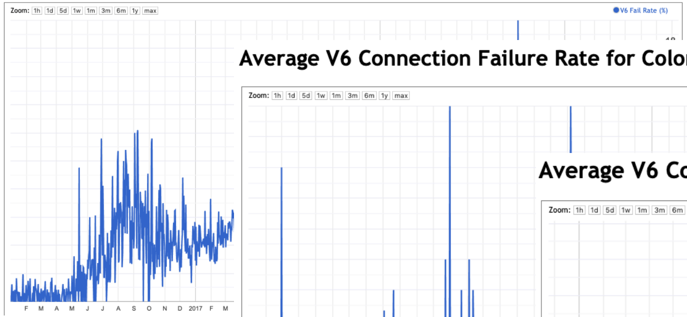


High IPv6 Failure Levels

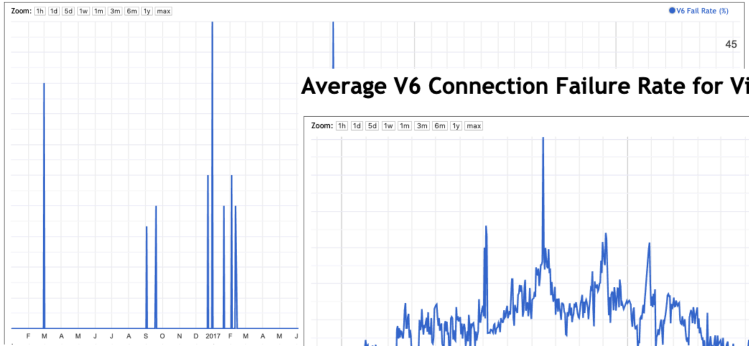
Average V6 Connection Failure Rate for Guatemala (GT)



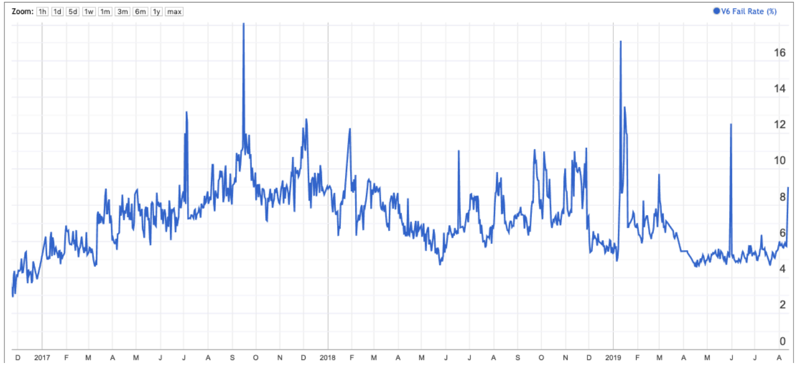
Average V6 Connection Failure Rate for New Zealand (NZ)



Average V6 Connection Failure Rate for Colombia (CO)

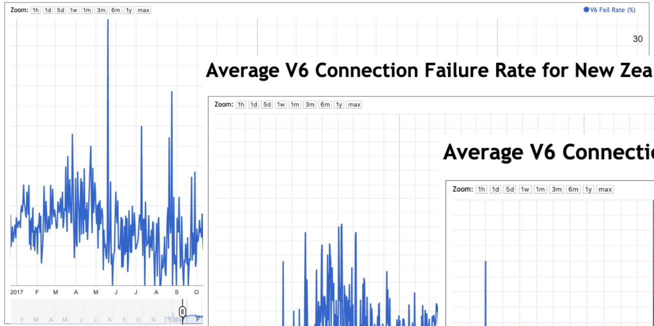


Average V6 Connection Failure Rate for Vietnam (VN)

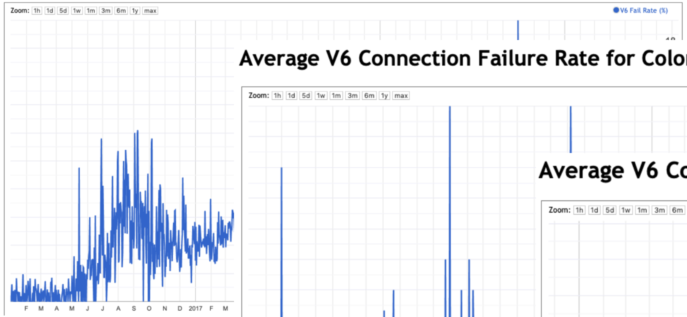


High IPv6 Failure Levels

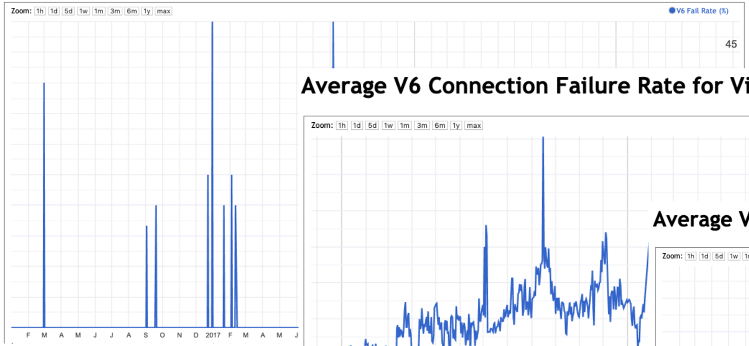
Average V6 Connection Failure Rate for Guatemala (GT)



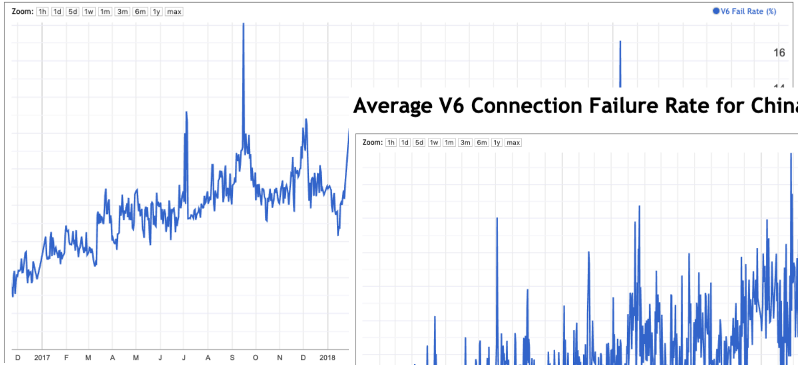
Average V6 Connection Failure Rate for New Zealand (NZ)



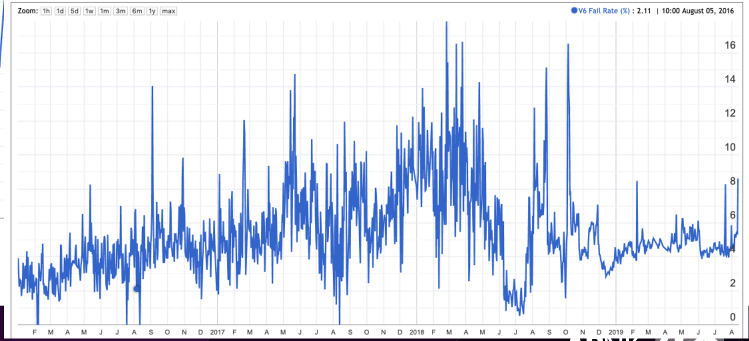
Average V6 Connection Failure Rate for Colombia (CO)



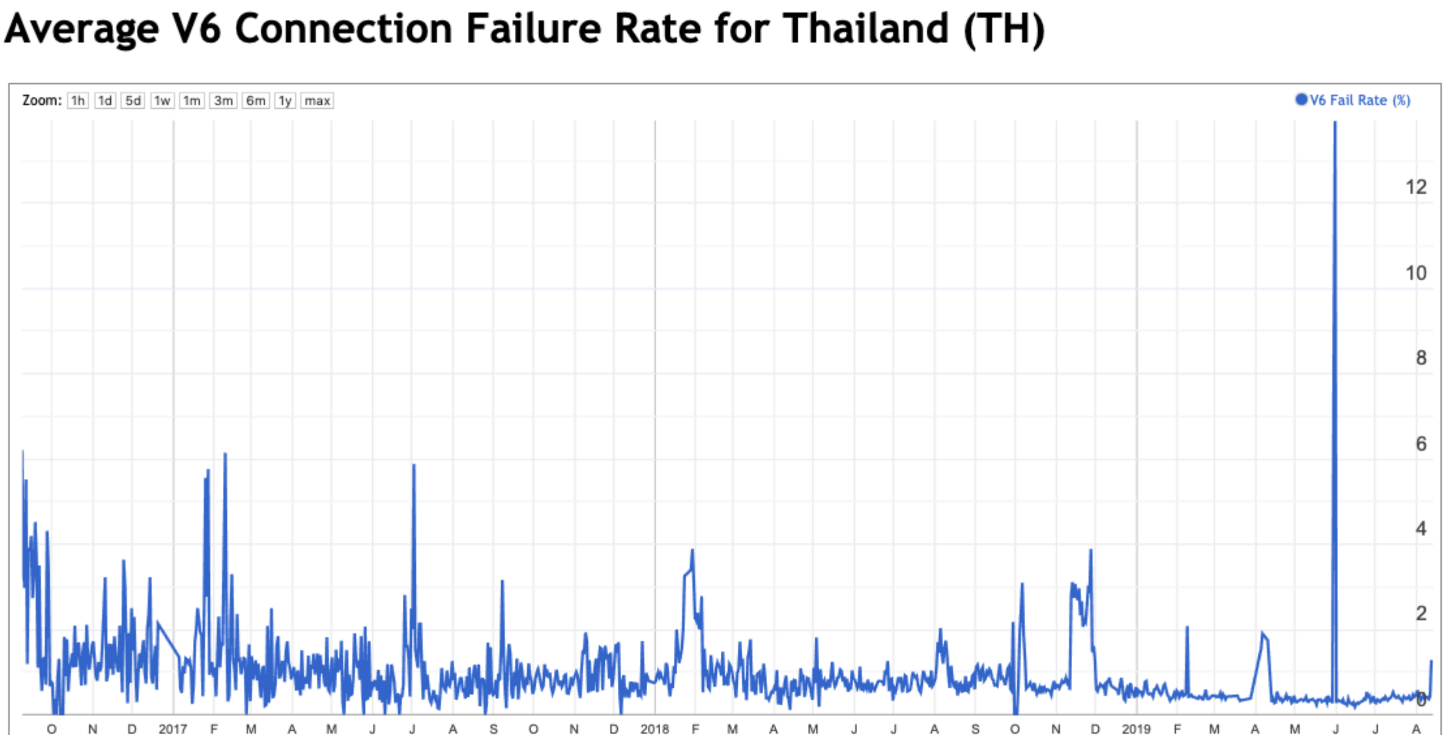
Average V6 Connection Failure Rate for Vietnam (VN)



Average V6 Connection Failure Rate for China (CN)

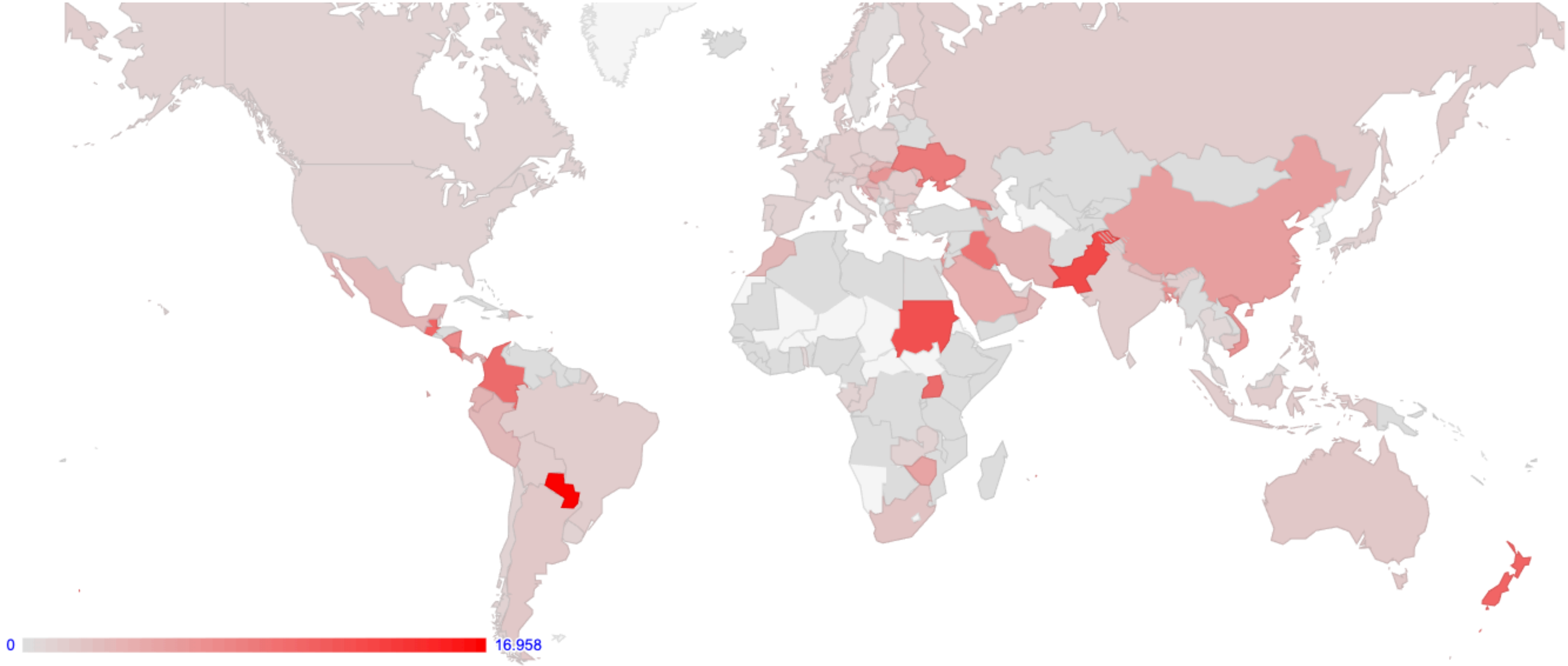


Thailand



Average V6 Connection Failure Rate for Thailand (TH)
Zoom: 1h 1d 5d 1w 1m 3m 6m 1y max
V6 Fail Rate (%)
12
10
8
6
4
2
0
O N D 2017 F M A M J J A S O N D 2018 F M A M J J A S O N D 2019 F M A M J J A

The big picture



APNIC 48

Comment

- For many end users in Vietnam, New Zealand, Guatamala, Colombia and China their IPv6 service looks pretty broken
  - The combination of Dual Stack and Happy Eyeballs masks the problem so that the user does not experience a degraded service
  - But this only will work while Dual Stack is around
- Other ISPs have managed to do a much better job, such as in the United States, Sweden, Thailand and Korea and the IPv6 connection failure rates are close to experimental noise levels
- What’s happening in the second set of countries and ISPs that is NOT happening in the first set?

Transition Technologies

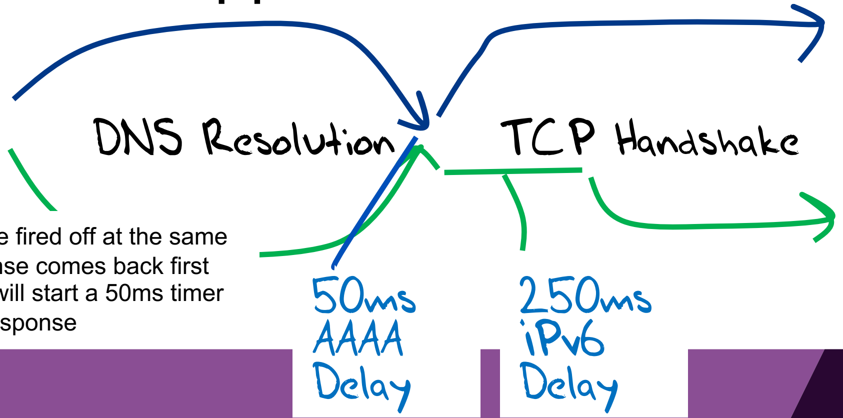
- Stateful transition technologies that involve protocol translation show higher levels of instability
- Translation technologies that require orchestration of DNS and network state are also more unstable

Dual Stack is NOT the Goal

- Despite all the grim predictions that IPv4 will be around for a long time to come, the aim of this transition is NOT to make Dual Stack work optimally
- The goal is to automatically transition the network to operate over IPv6
- The way to achieve this is for client systems to prefer to use IPv6 whenever it can

Happy Eyeballs

- An unconditional preference for IPv6 can lead to some very poor user experience instances
  - Linux uses a 108 second connection timer, for example
- Applications (particularly browsers) have used a “Happy Eyeballs” approach



DNS A and AAAA are fired off at the same time – if the A response comes back first then the application will start a 50ms timer to wait for a AAAA response

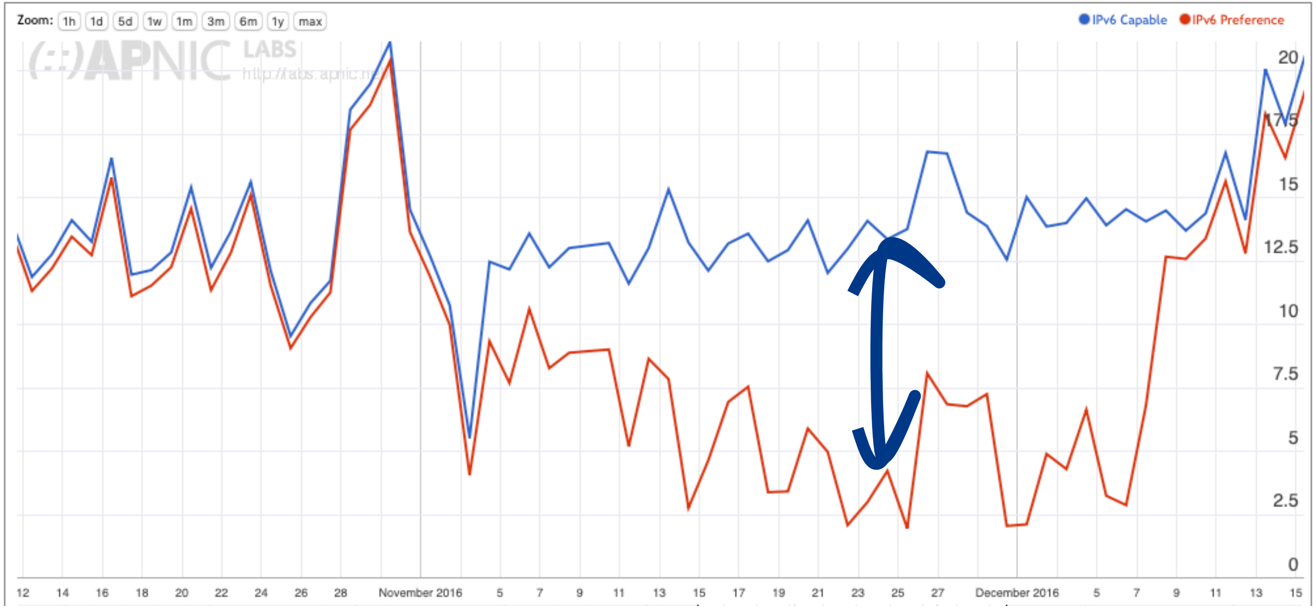
A TCP session will be started in IPv6 if there is a IPv6 address record. If the handshake is not completed within 250 ms then an IPv4 TCP session is also fired off

Tuning IPv6 for Happy Eyeballs

- When connecting to a remote dual stack service, the Routing Path selection for IPv6 should be similar to IPv4
- Where there are path deviations, the path discrepancy should be contained
- This is not always the case…

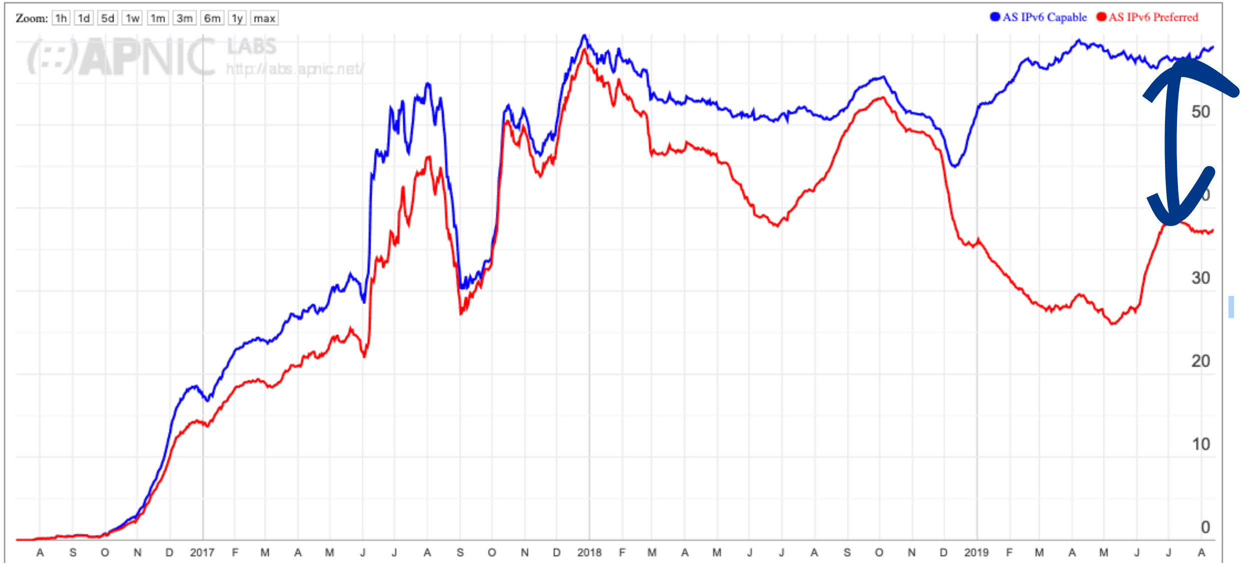
India, late 2016

Use of IPv6 for India (IN)

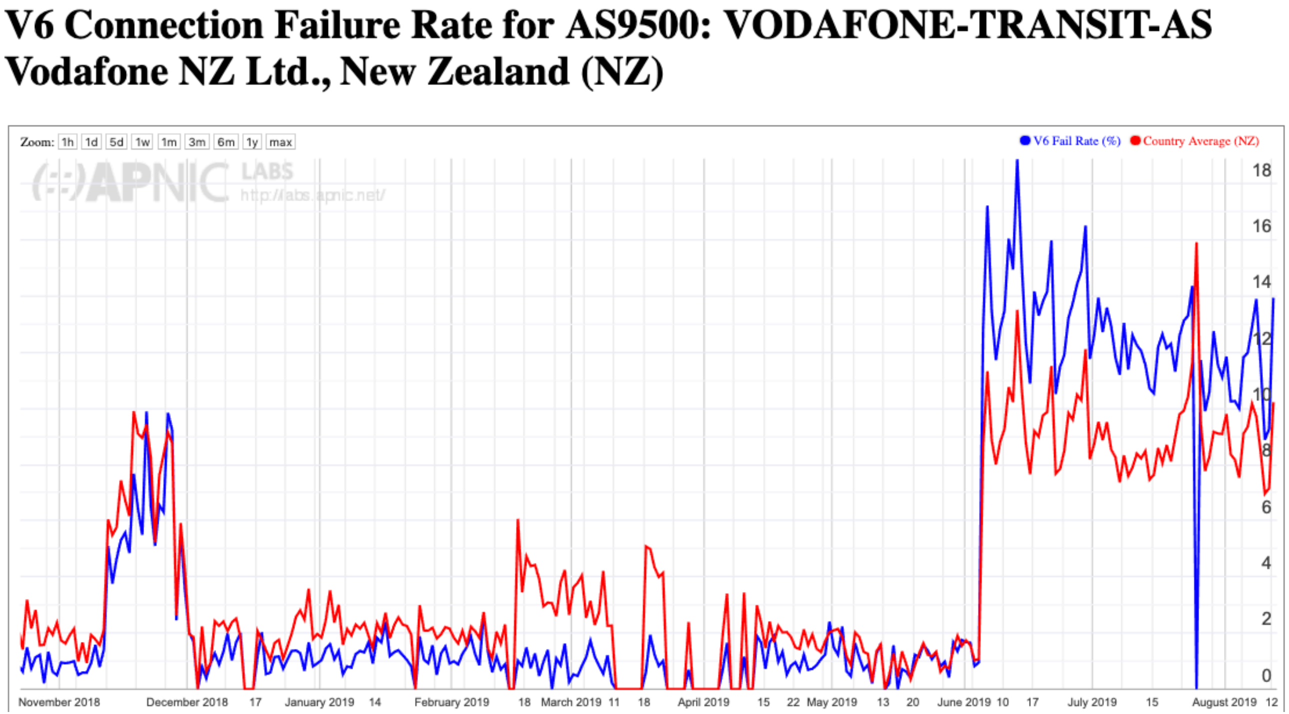


Vodaphone New Zealand - 2019

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



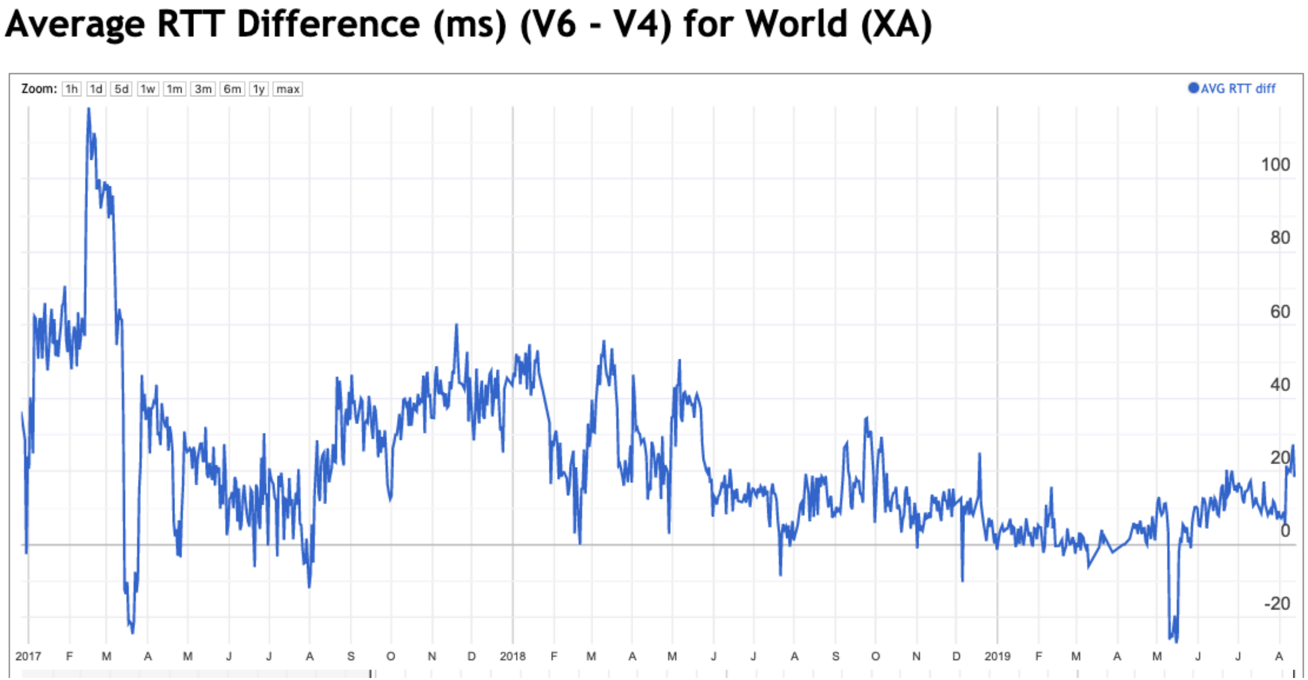
Vodaphone New Zealand - 2019



Sometimes it's the DNS!

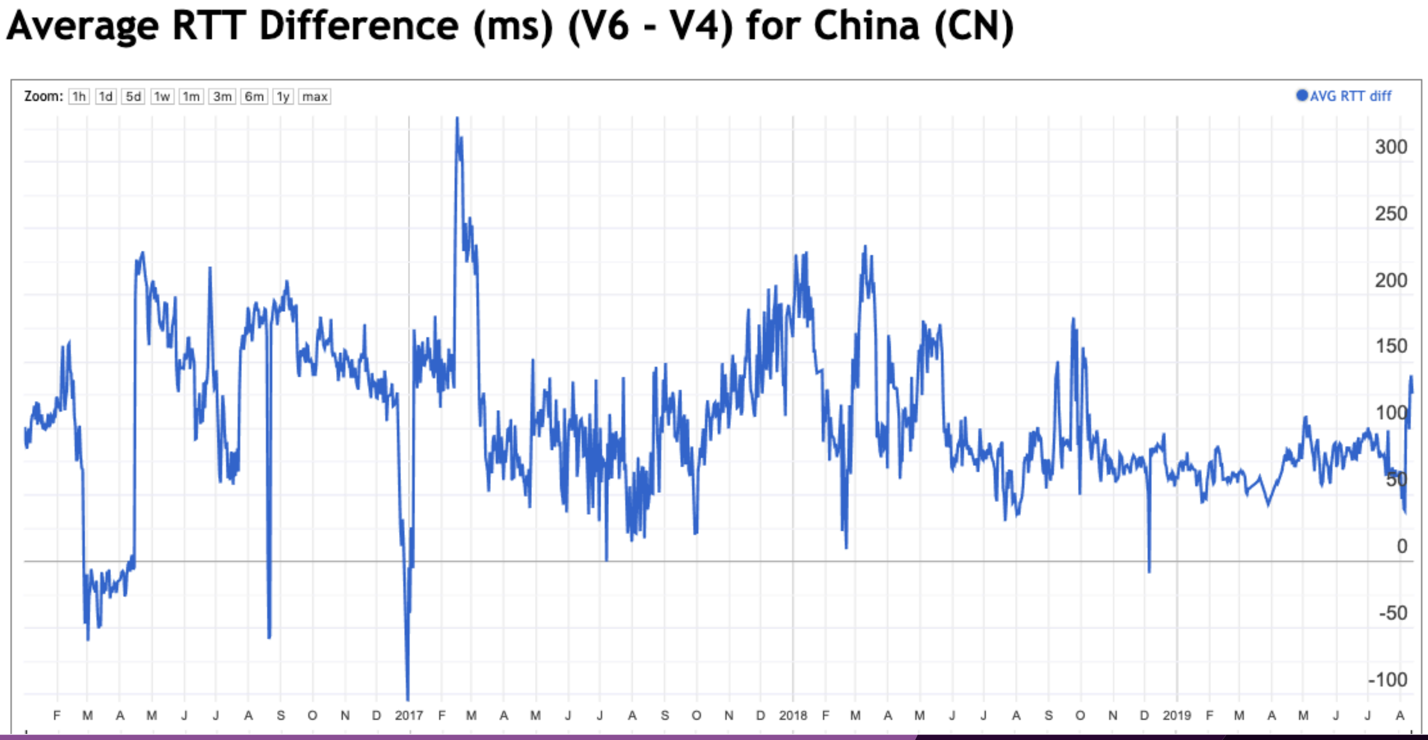
- Happy Eyeballs assumes that the time to resolve an A and a AAAA record are within 50 msecs of each other
- The client generates a query for the A record and a second query for a AAAA record at the same time
- The recursive resolver does not necessarily process the two requests in parallel:
  - A QNAME minimisation resolver may use A queries to walk the DNS hierarchy
  - A DNS-based content filter may use A queries to determine the outcome
  - In a partial deployed state its more likely that the A record is already in the local cache

Worldwide RTT Diff Performance



The overall trend was improving through 2018. The picture appears to have reversed in 2019

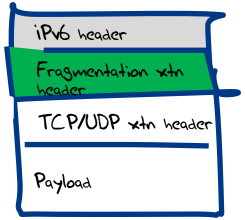
China's IPv6 Network



3 Suggestions to Assist IPv6 Robustness

- Avoid stateful IPv6 -> IPv4 transition mechanisms if possible – if you can operate IPv6 in native mode all the better!
- Avoid using IPv6-in-IPv4 encapsulations
  - Not only are tunnels unstable, but the reduced IPv6 MTU may cause problems with extension header based packet discard
- Keep IPv4 and IPv6 paths congruent if possible
  - Yes, this can be really challenging for multi-homed networks, but try to use transit and peer arrangements that are dual stack

But that's not all…



- IPv6 used a new approach to extension headers, including packet fragmentation by inserting them between the IPv6 header and the transport header
- Which means that hardware will have to spend cycles to hunt for a transport header
- Or it can just drop the packet…

2017 Measurement

V6, the DNS and Fragmented UDP

Total number of tests: 10,851,323

Failure Rate in receiving a large response: 4,064,356

IPv6 Fragmentation Failure Rate: **38%**

This measurement test involved sending a fragmented UDP packet to recursive resolvers

2017 Measurement

What about TCP and Fragmentation?

1,961,561 distinct IPv6 end point addresses

434,971 failed to receive Fragmented IPv6 packets

22% failure rate

This measurement test involved sending a fragmented TCP packet to browser endpoints

What can we say?

- There are ongoing issues with IPv6 reliability in many parts of the world
  - This appears to relate to local security policies at the client edge of the network
  - We can expect most of this to improve over time by itself

What can we say?

- But there are also very serious issues with Path MTU management and handling of IPv6 extension headers
  - This is a more challenging issue that will probably not just clean itself up over time

  - Should we just avoid IPv6 extension headers?
  - Or try to clean up the IPv6 switching infrastructure?

What can we say?

- But there are also very serious issues with Path MTU management and handling of IPv6 extension headers
  - This is a more challenging issue that will probably not just clean itself up over time

  - Should we just avoid IPv6 extension headers?
  - ~~Or try to clean up the IPv6 switching infrastructure?~~ Unlikely!

Thanks!