IPv6 Performance Measurement

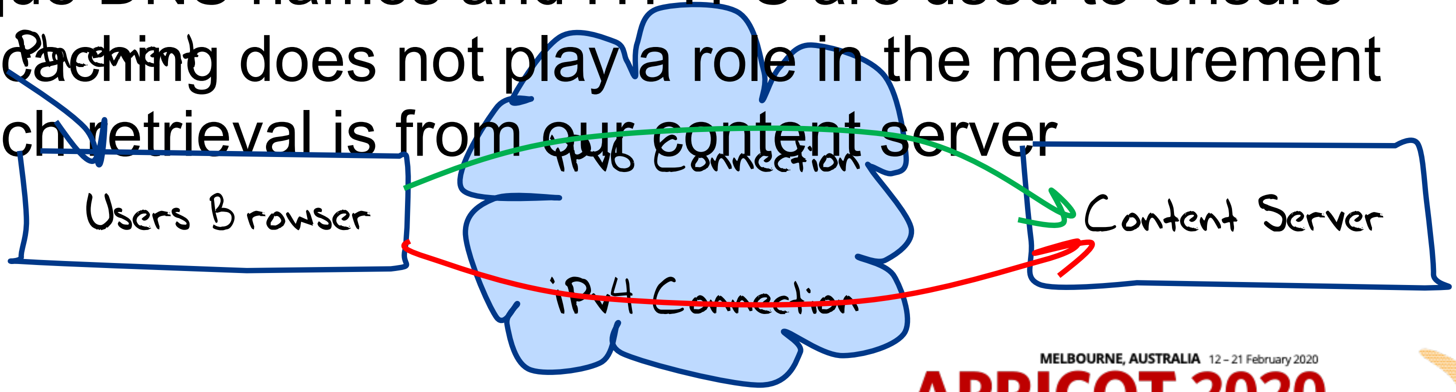
February 2020

Geoff Huston
APNIC

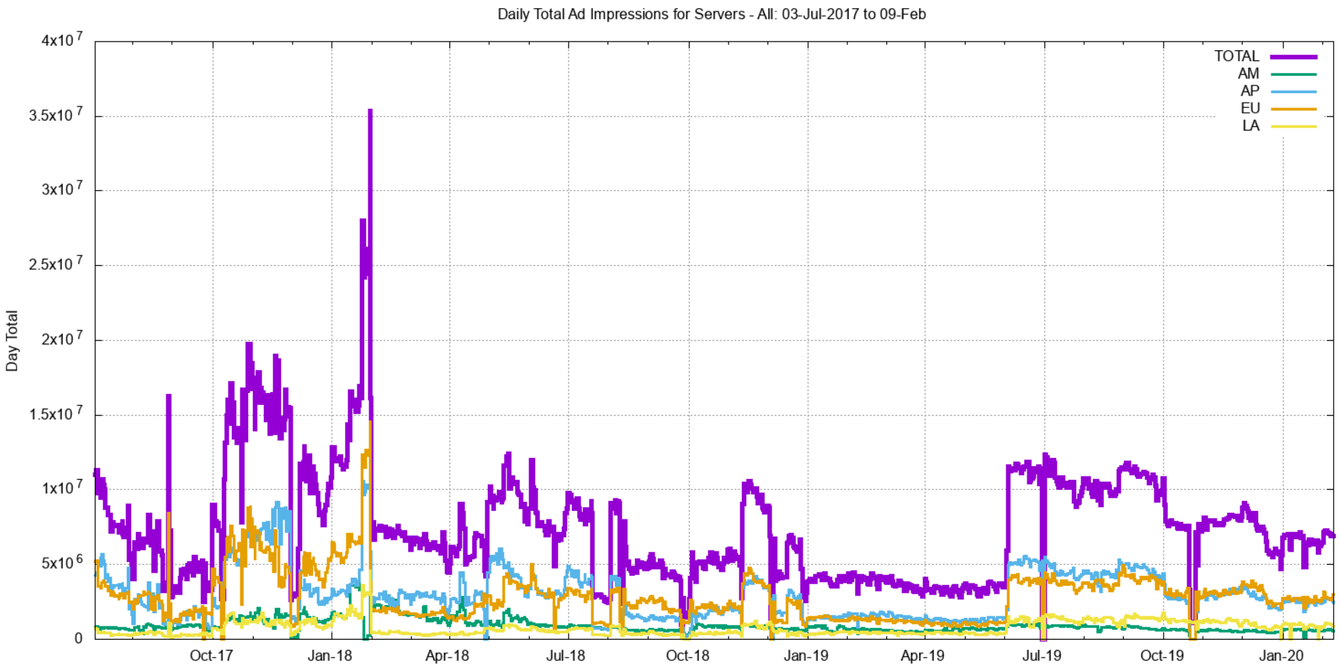
The Measurement

The endpoint that runs the experiment attempts to retrieve two URLs from the same remote server – one using IPv4 and the other using IPv6

– Unique DNS names and HTTPS are used to ensure that caching does not play a role in the measurement – each retrieval is from our content server



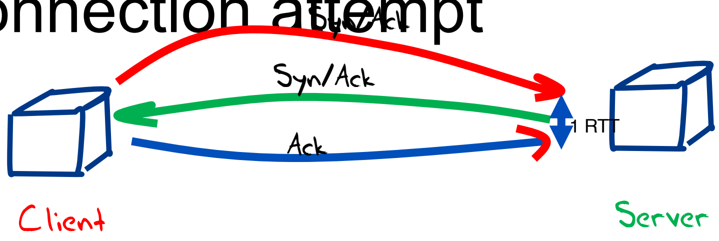
Measurement Volume



Ad Impressions per day for the period 2017 - 2020

The Measurement

- We perform full packet capture at the server
- Data analysis
  - We look at the SYN/ACK exchange at the start of the TCP session
  - A received SYN with no subsequent ACK is interpreted as a failed connection attempt



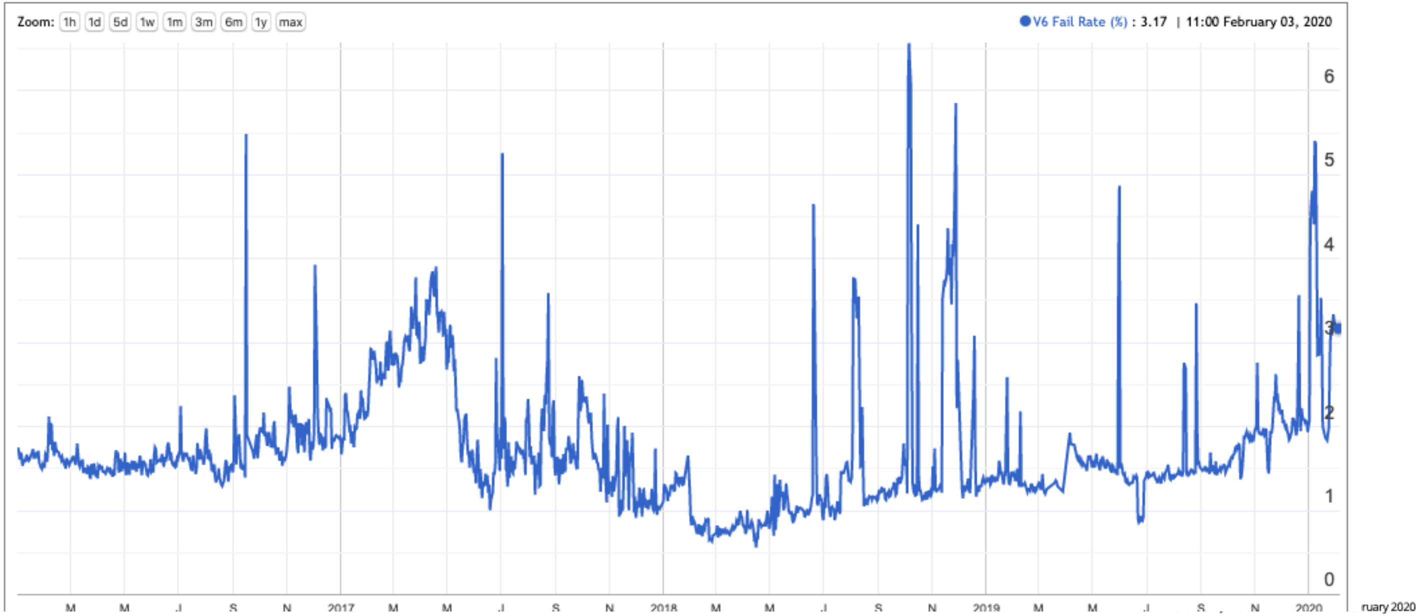
Analysis - Reliability

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

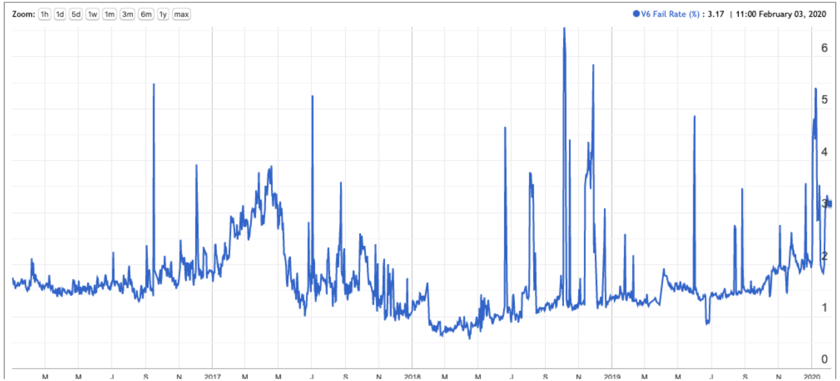
IPv6 TCP Connection Failure

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 2-3% is getting worse! As the IPv6 network is growing, its performance in terms of reliability is getting worse

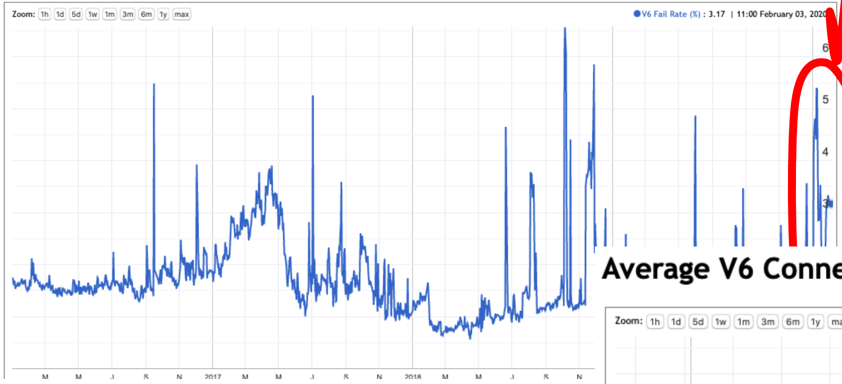
What we are seeing is most likely a failure to deliver an IPv6 packet from the server to the endpoint

Possible reasons:

- Endpoint using an unreachable IPv6 address
- End site firewalls
- ??

IPv6 TCP Connection Failure

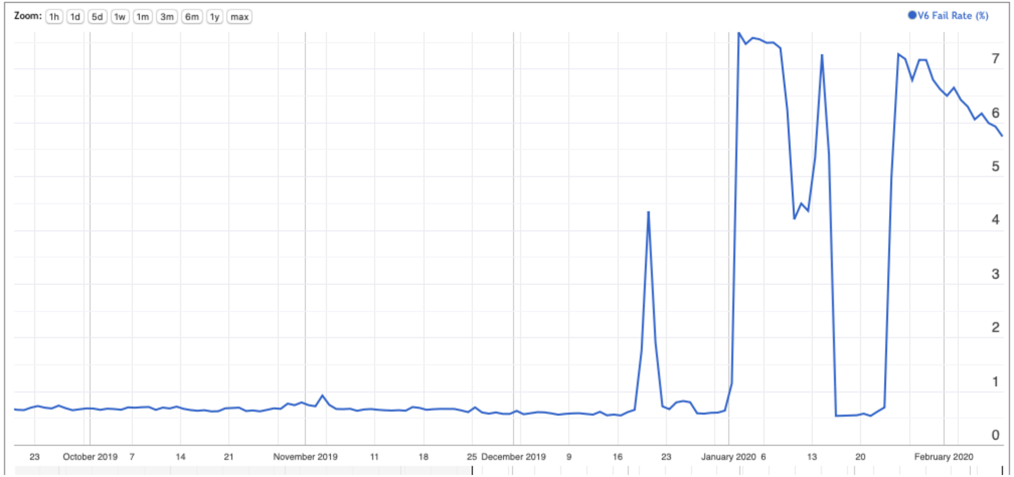
Average V6 Connection Failure Rate for World (XA)



There have been some recent high noise periods

This is due to IPv6 routing instability in the North American network – parts of the IPv6 routing table appear to have been dropped for some destinations

Average V6 Connection Failure Rate for Northern America (XQ)



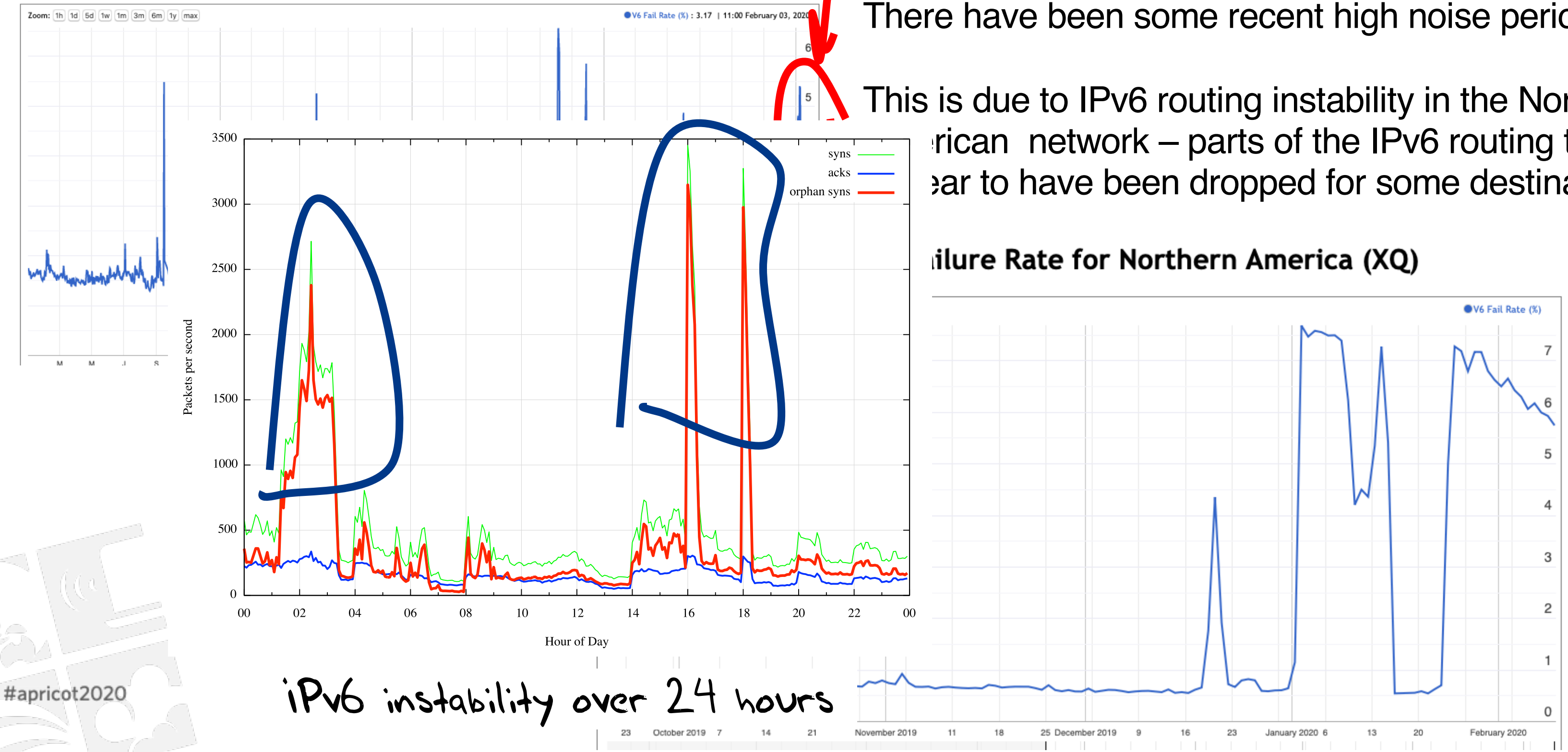
IPv6 TCP Connection Failure

Average V6 Connection Failure Rate for World (XA)

There have been some recent high noise periods

This is due to IPv6 routing instability in the North
...rican network – parts of the IPv6 routing table
...ear to have been dropped for some destinations

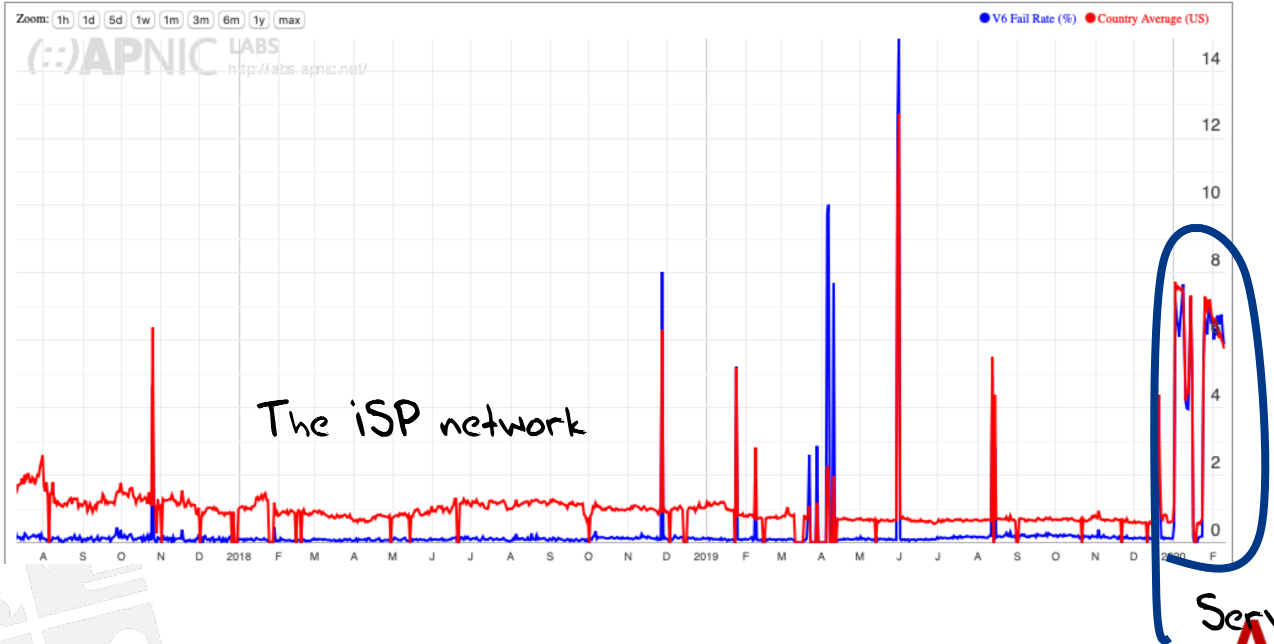
...ilure Rate for Northern America (XQ)



iPv6 instability over 24 hours

The Good

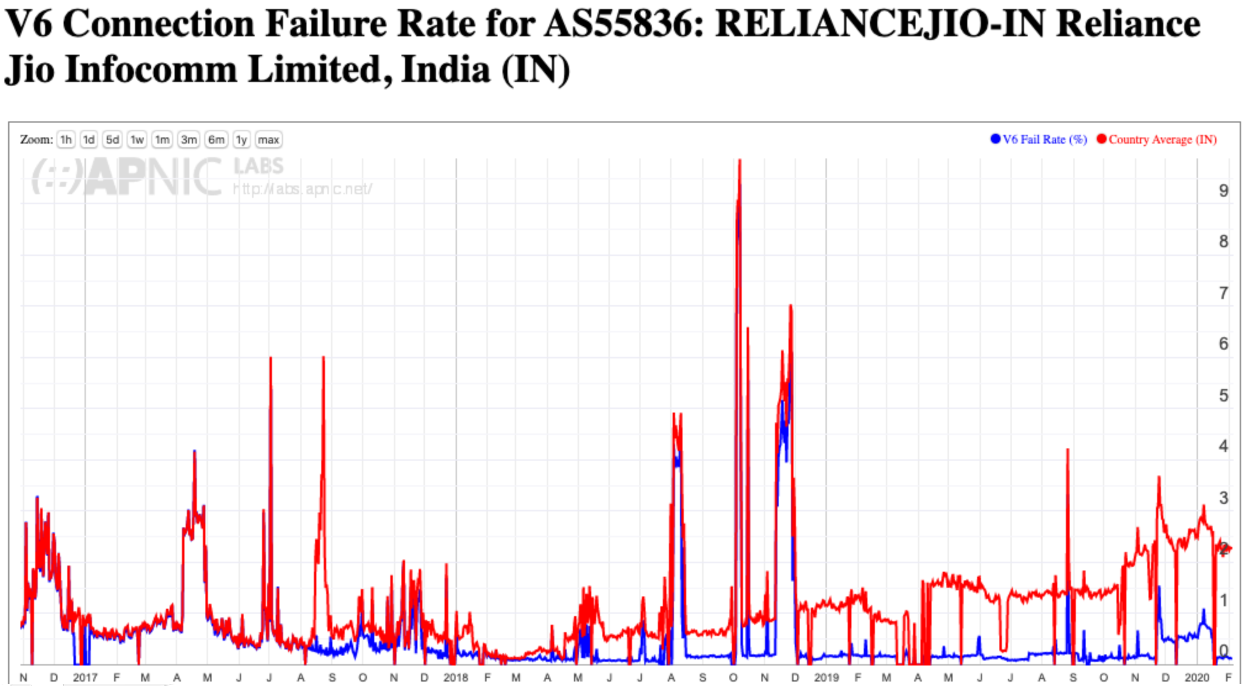
V6 Connection Failure Rate for AS21928: T-MOBILE-AS21928, United States of America (US)



Server Platform problems

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

The Good



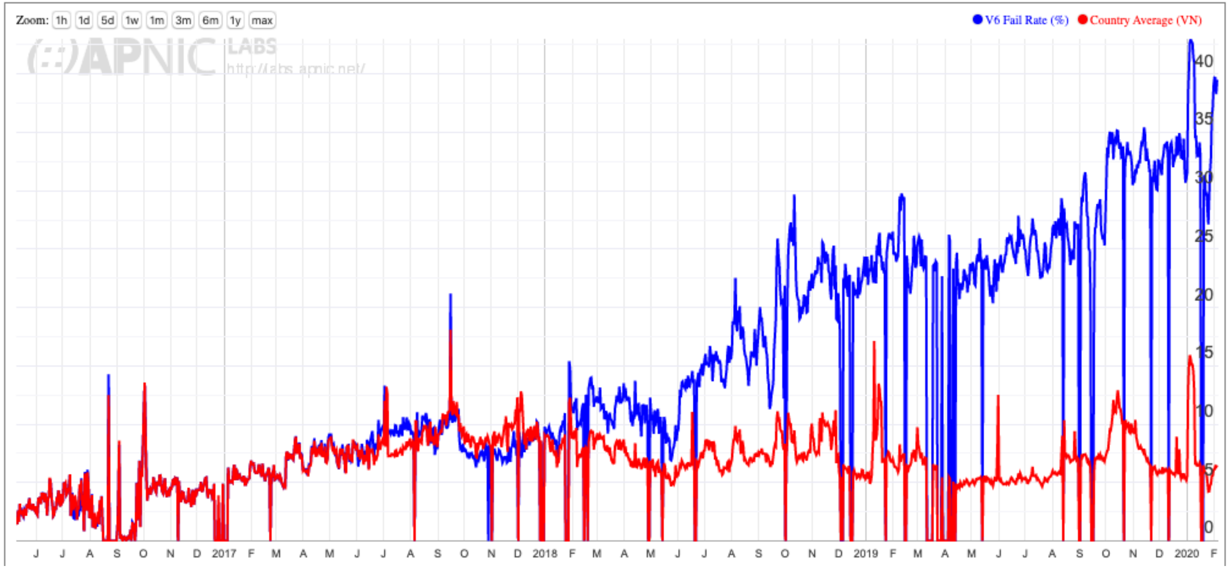
Similar story in India with Reliance JIO – the endpoints are connected via native IPv6 and as this is a mobile network there is only a small amount of customer-operated filtering middleware

464XLAT Performance

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

The not quite so good

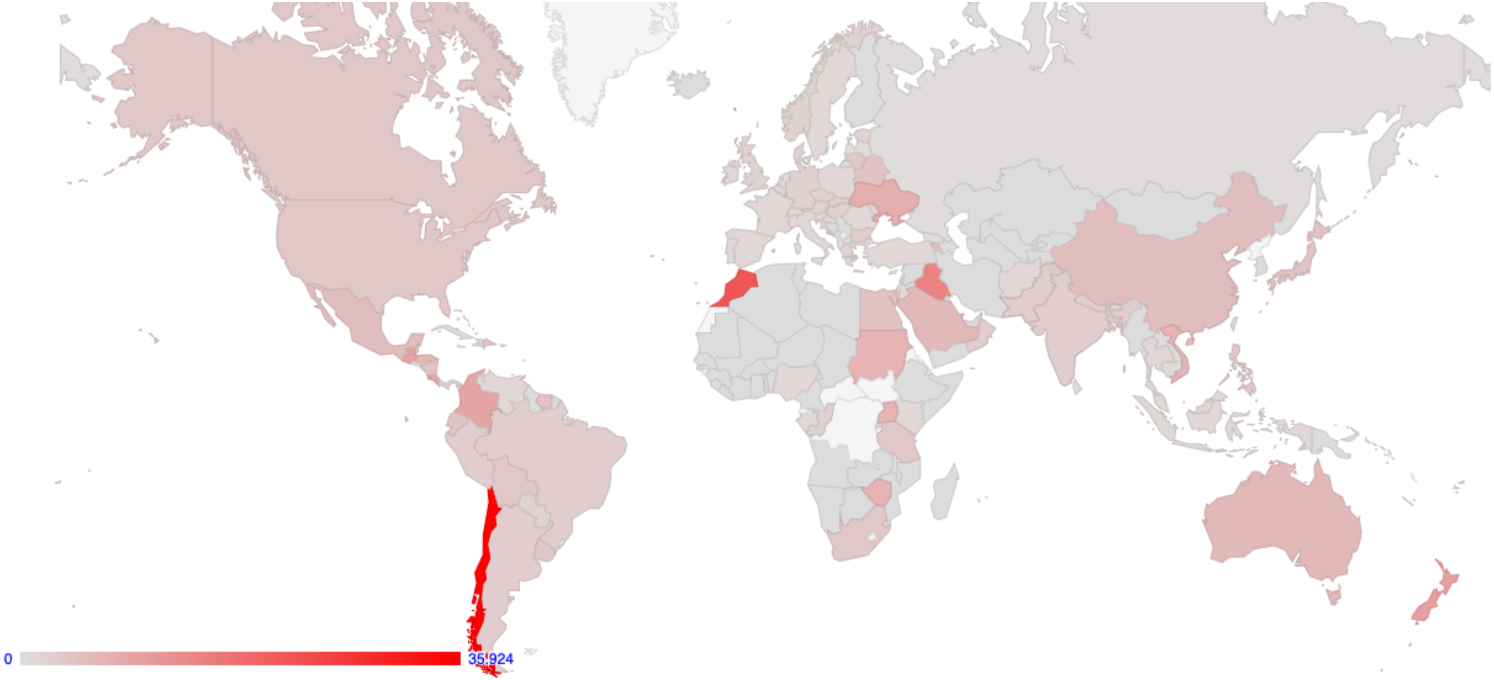
V6 Connection Failure Rate for AS18403: FPT-AS-AP The Corporation for Financing & Promoting Technology, Vietnam (VN)



January 2020 Stats

Code	Country	Avg RTT Diff (V6-V4)	Samples	Avg V6 Fail Rate	V6 Fails	V6 Samples	Dual Stack	Dual Stack (300ms)	V6 Use Rate
CL	Chile, South America, Americas	2.34 ms	708	33.14%	395	1,192	86.72%	90.25%	0.12%
CR	Costa Rica, Central America, Americas	-4.64 ms	219	13.27%	41	309	79.00%	85.39%	0.14%
MA	Morocco, Northern Africa, Africa	-24.73 ms	54	11.67%	7	60	90.74%	100.00%	0.01%
CO	Colombia, South America, Americas	7.67 ms	47,310	11.29%	8,084	71,578	48.54%	98.34%	2.57%
IQ	Iraq, Western Asia, Asia	-4.82 ms	36	11.11%	4	36	100.00%	100.00%	0.00%
FO	Faeroe Islands, Northern Europe, Europe	-11.23 ms	135	10.70%	20	187	80.74%	99.26%	6.36%
NZ	New Zealand, Australia and New Zealand, Oceania	-32.56 ms	17,722	10.64%	2,651	24,907	68.45%	87.08%	24.84%
GT	Guatemala, Central America, Americas	1.64 ms	22,598	10.27%	3,254	31,697	31.93%	99.19%	11.30%
SD	Sudan, Northern Africa, Africa	15.58 ms	33	10.00%	5	50	57.58%	100.00%	0.04%
UA	Ukraine, Eastern Europe, Europe	-2.11 ms	5,436	8.64%	557	6,447	65.18%	95.81%	0.24%
UG	Uganda, Eastern Africa, Africa	-4.81 ms	112	7.64%	11	144	33.04%	94.64%	0.17%
AM	Armenia, Western Asia, Asia	-6.88 ms	6,520	7.43%	664	8,941	66.26%	99.17%	7.13%
QA	Qatar, Western Asia, Asia	120.15 ms	46	7.02%	4	57	13.04%	97.83%	0.03%
ZW	Zimbabwe, Eastern Africa, Africa	-13.01 ms	4,058	7.00%	390	5,575	75.18%	89.35%	10.47%
VN	Vietnam, South-Eastern Asia, Asia	-6.35 ms	999,609	6.74%	90,799	1,346,562	46.72%	98.45%	42.74%
SX	Sint Maarten (Dutch part), Caribbean, Americas	-53.84 ms	15	6.45%	22	341	53.33%	100.00%	0.47%
TT	Trinidad and Tobago, Caribbean, Americas	-32.94 ms	15,603	6.33%	1,243	19,627	89.08%	99.15%	22.72%
MX	Mexico, Central America, Americas	-33.72 ms	1,104,014	6.11%	90,783	1,485,825	81.91%	99.03%	32.69%
SA	Saudi Arabia, Western Asia, Asia	-20.90 ms	124,098	5.64%	9,697	171,984	89.34%	98.03%	13.39%
CN	China, Eastern Asia, Asia	82.75 ms	474,004	5.45%	38,838	713,271	42.07%	84.75%	16.36%
EG	Egypt, Northern Africa, Africa	-39.85 ms	136,429	5.36%	9,799	182,723	82.36%	99.07%	11.14%
BY	Belarus, Eastern Europe, Europe	-1.54 ms	91	5.32%	5	94	40.66%	95.60%	0.02%

The Bigger Picture of IPv6 Connection Failure



Comment

- For many end-users 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 India, Iceland, Australia and Korea and the IPv6 connection failure rates are close to experimental noise levels

Transition Technologies

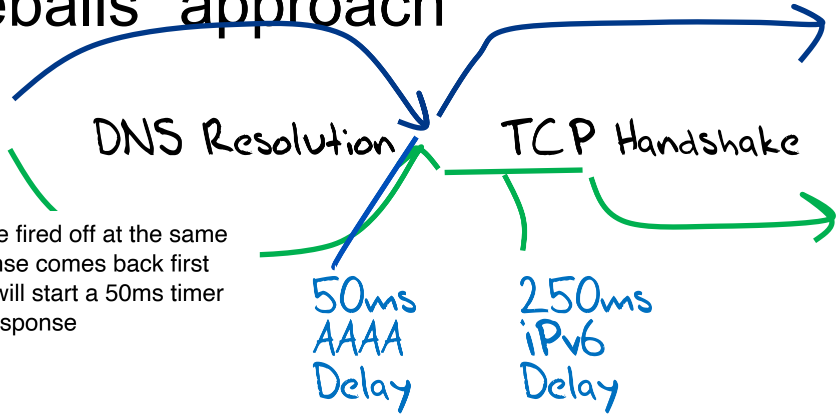
- Stateful transition technologies that involve protocol translation show higher levels of instability
- Translation technologies that require orchestration of DNS and network state are also more unstable

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



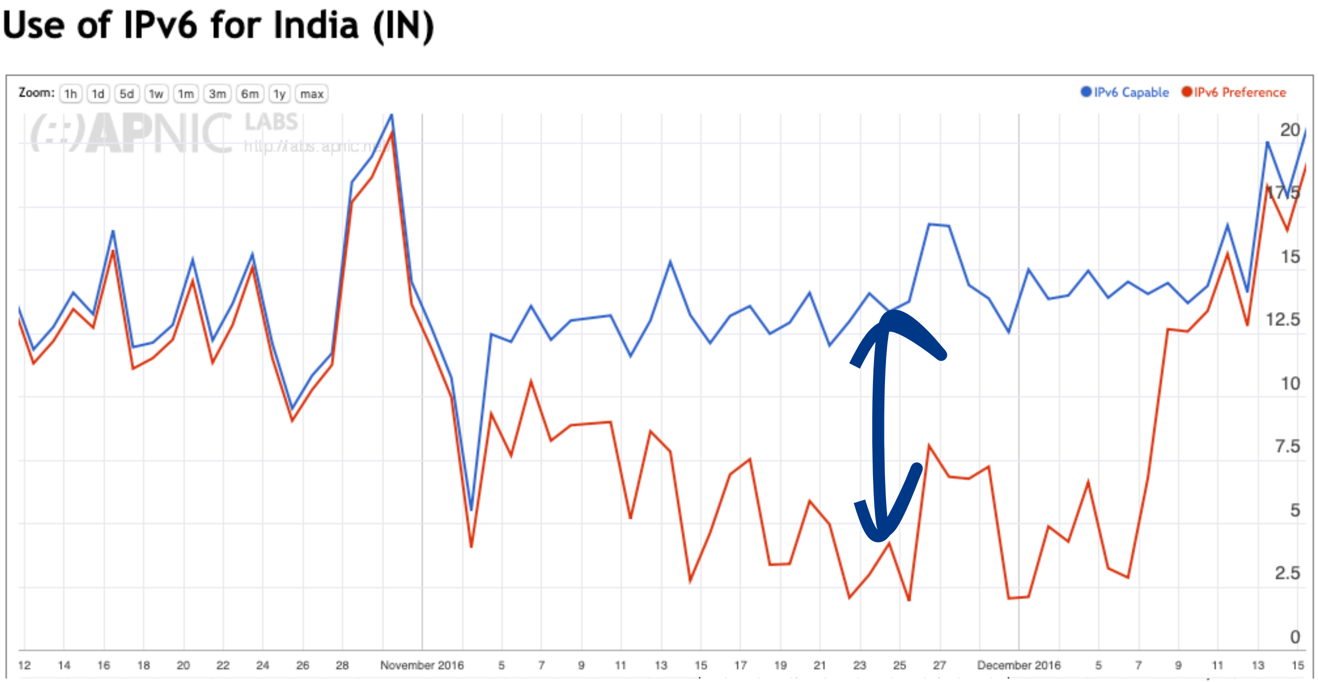
DNS A and AAAA are fired off at the same time – if the A response comes back first then the application will start a 50ms timer to wait for a AAAA response

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

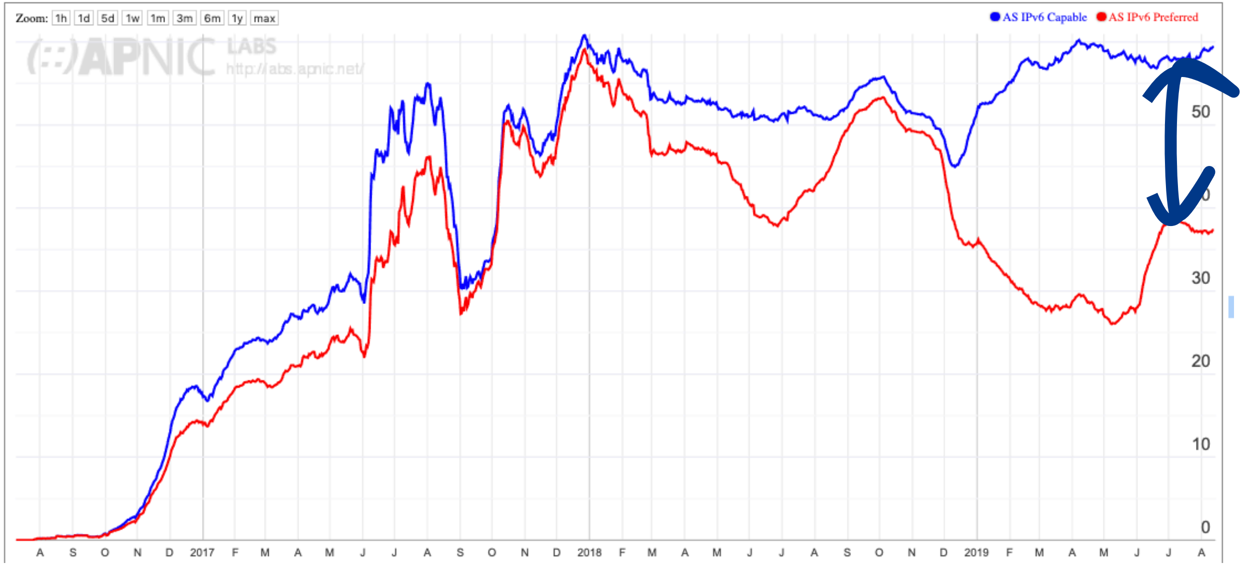
- When connecting to a remote dual stack service, the Routing Path selection for IPv6 should be similar to IPv4
- Where there are path deviations, the path discrepancy should be contained
- This is not always the case…

India, late 2016



Vodaphone New Zealand – 2019

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



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

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

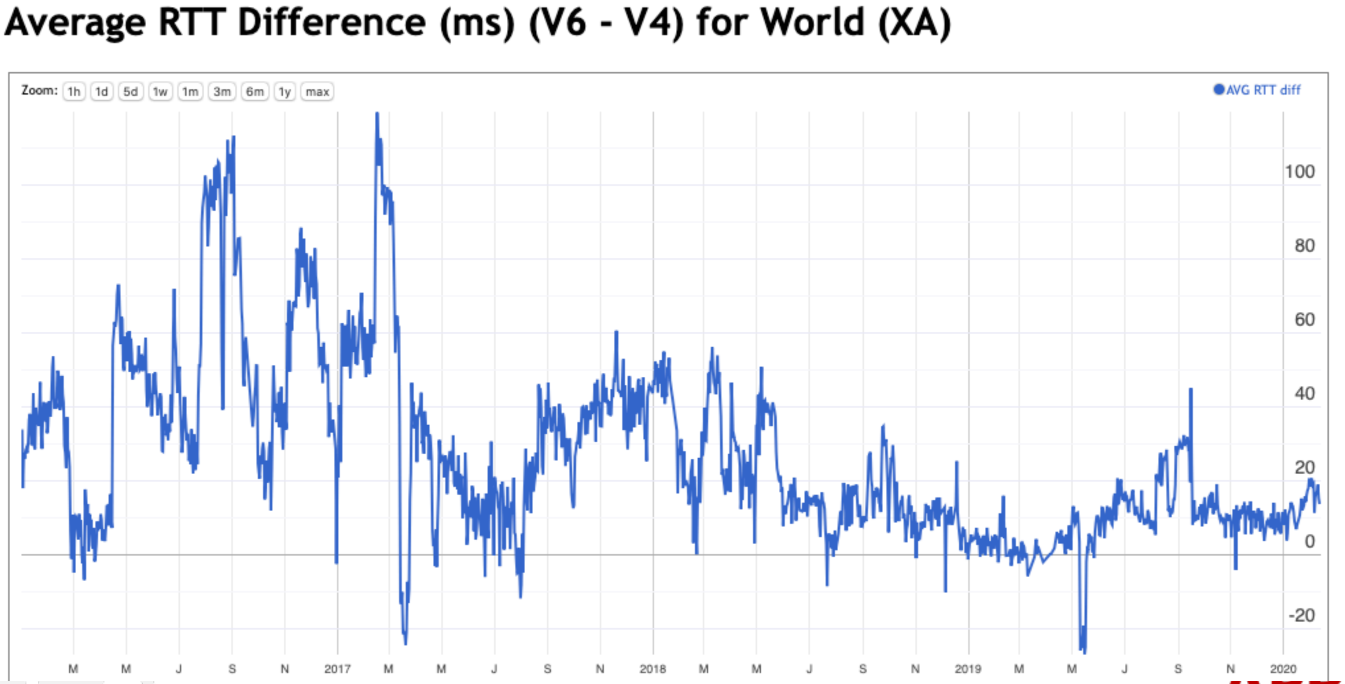
Speed Measurement

- We perform full packet capture at the server
- Data analysis
  - We look at the SYN/ACK exchange at the start of the TCP 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)

Analysis - Speed

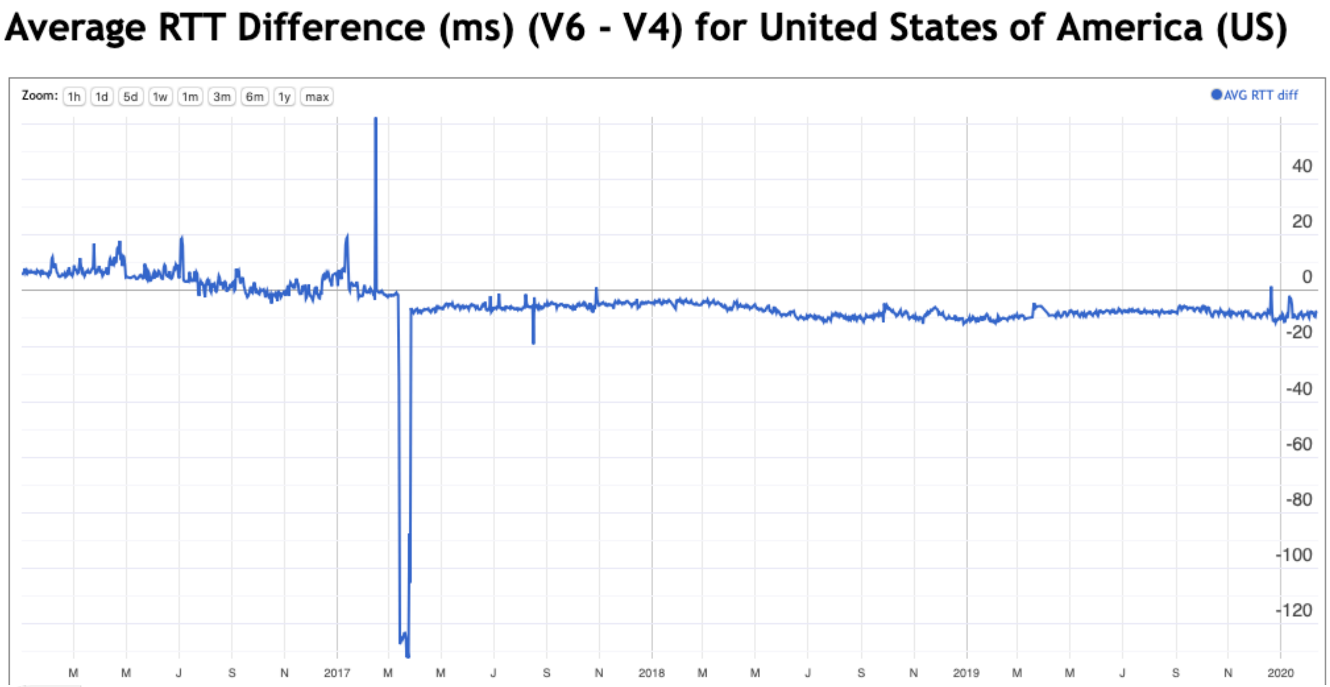
- Why measure only the handshake delay? Why not measure a larger data transfer?
- Because in the end host and the server the same TCP version is used on top of IPv4 and IPv6
  - If the end to end paths are the same in IPv4 and IPv6 we would see precisely the same session throughput
- RTT and packet loss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

Worldwide RTT Diff Performance



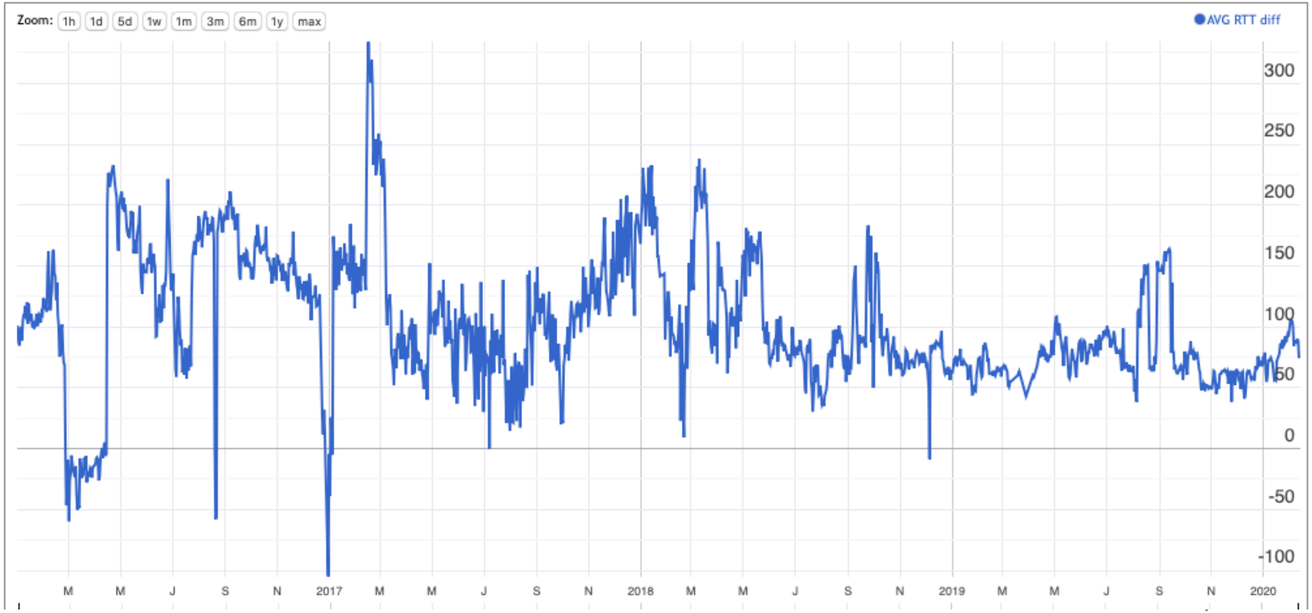
IPv4 is consistently faster than IPv6 on average

US IPv6 Network

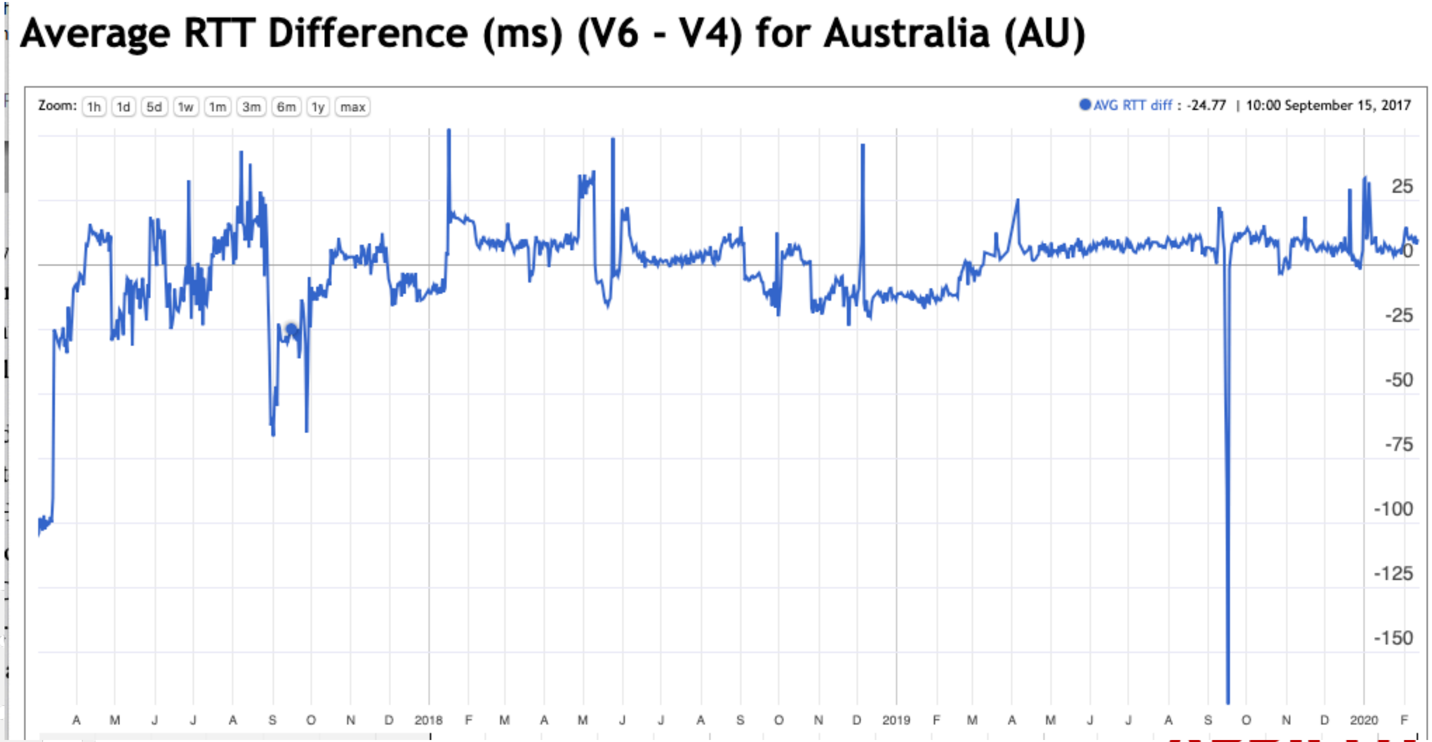


China's IPv6 Network

Average RTT Difference (ms) (V6 - V4) for China (CN)



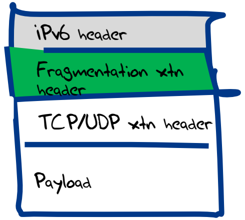
Australia's IPv6 Network



This is a localised measurement

- This is the result of millions of endpoints heading to one of 4 measurement points
  - If IPv4 and IPv6 paths are aligned then the RTT diff would be close to zero
  - Any deviation points to some form of asymmetric routing issues
  - And whether IPv6 is faster or slower than IPv4 is less important than the fact that they are different
- But the observation that they are different with respect to a

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

Unlikely!

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

Thanks!

25 YEARS

2020 APRICOT

APNIC 49

MELBOURNE
AUSTRALIA

12 – 21 February 2020

#apricot2020