A Longitudinal View of Netflix: Content Delivery over IPv6 and Content Cache Deployments

Trinh Viet Doan*, Vaibhav Bajpai*, Sam Crawford†

*Technical University of Munich (TUM),
†SamKnows

IEEE INFOCOM,
July 6–9, 2020

Motivation

Motivation

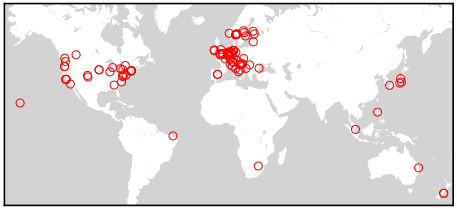
- Video streaming accounts for majority of Internet traffic
  - YouTube 8.7%, Netflix 12.6% of all global downstream traffic[1] as of 2019
- Introduction of Netflix Open Connect in 2011
- Open Connect Appliances (OCAs) deployed within IXPs and ISPs to bring content to the edge
- YouTube and Netflix as main drivers of IPv6
  - 31.5% of Google users with available IPv6 connectivity[2] as of 2020

$\Rightarrow$ Longitudinal measurement study on Netflix OCAs from multiple vantage points with focus on content delivery over IPv6 and content caches

[1] Sandvine, Global Internet Phenomena Report 2019, `https://bit.ly/3cvN5Qi`
[2] Google IPv6 Statistics, `https://www.google.com/intl/en/ipv6/statistics.html`

Methodology

Methodology



Source:
https://samknows.one/hc/en-gb/articles/360000451757

- Roughly 100 probes deployed in 74 origin ASes (primarily home networks)
- Dual-stack with native IPv6 connectivity, i.e., no transition mechanisms at access network
- Measurement period:
  July 2016 – April 2019

- Hardware:
  dedicated SamKnows Whiteboxes (formerly TP-Link routers with customized OpenWrt)
- Running set of measurement tests periodically

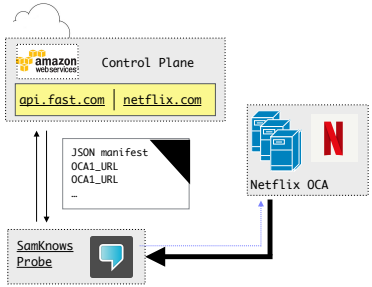
Methodology

Tests:

- **`netflix`** test and **`traceroute`** (`scamper`) to Netflix OCAs
- **`speedtest`** to Measurement Lab (M-Lab) servers
- Schedule: tests run every hour over IPv4 and IPv6 each ("pair")

`netflix` test:

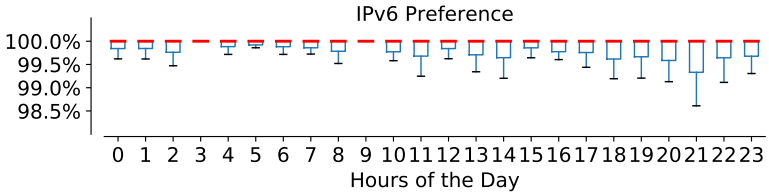
1. Authenticate with control plane
2. Receive list of OCAs determined by `fast.com` API
3. Connect to OCA
4. (Repeatedly) download 25 MB file for 20 sec to simulate streaming



Analysis: Content Delivery

- How does Netflix content delivery perform w.r.t. latency and throughput over both address families?
- Do users benefit or suffer from downloading Netflix content over IPv6 compared to IPv4?

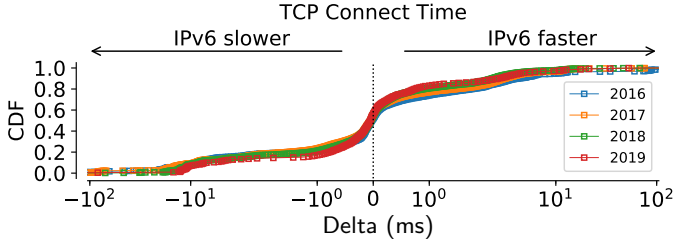
IPv6 Preference



- Calculate IPv6 preference based on Happy Eyeballs [RFC 8305]
- Group measurement pairs by local time
- Overall, very high IPv6 preference, i.e., IPv6 slower by at most 250 ms
- However, IPv6 connections preferred less during peak hours (18:xx–23:xx)

IPv6 preferred in nearly all cases when connecting to Netflix OCAs, although preference slightly drops during peak hours.

TCP Connect Time

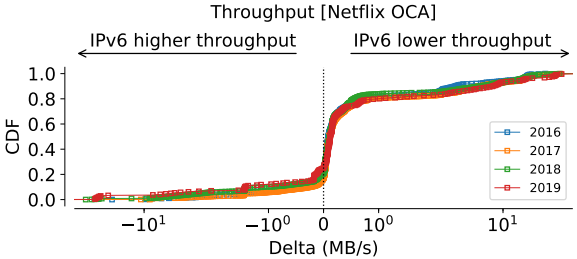


- Median TCP connect time deltas ($tconn_{IPv4} - tconn_{IPv6}$) of probes per day
- Most differences within $\pm$10 ms, converging toward 0 ms
- Across all years, 45%–48% of samples faster over IPv6
- Absolute connect times improved by 40% over the years (25 ms $\rightarrow$ 15 ms)

TCP connect times over IPv4 and IPv6 converge to similar values as of 2019, with roughly 50% of the samples being faster/slower over IPv6.

Throughput

- Median throughput deltas ($tp_{IPv4} - tp_{IPv6}$) of probes per day
- IPv6 higher throughput than IPv4 in only 17%–25% of samples



- However, 70–75% of samples within ± 1 MB/s
- Achieved throughput increases over the years, though not specific to Netflix
  - Similar observations for `speedtest` measurements toward M-Lab servers

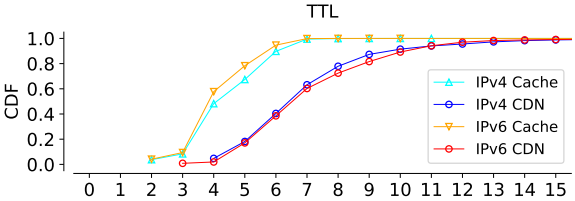
Although most of the deltas are within ± 1 MB/s, IPv6 shows lower throughput than IPv4 in 75–83% of the samples (not specific to Netflix).

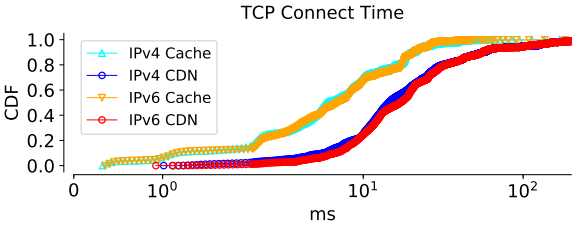
Analysis: Content Caches

- How do IP path lengths and latency compare between ISP caches and deployments outside the ISP boundary?
  - How do path lengths and latency differ over IPv4 and IPv6?
- How do content caches at the edge benefit content delivery?
  - How do these benefits compare over IPv4 and IPv6?

Content Caches

- Comparison of ISP caches and CDN servers
- Measurement distribution similar over IPv4 and IPv6
- 90% of ISP caches within 6 IP hops & 21 ms
- 90% of CDN servers within 10–12 IP hops & 56–58 ms
- Higher throughput with caches: 10–11 → 32–34 MB/s

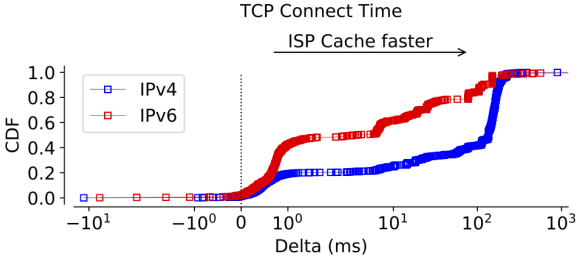




ISP caches are reachable within 6 IP hops and roughly 20 ms;
IP path lengths are shorter by 40–50%, TCP connect times lower by 64%,
throughput higher by factor of three.

Content Caches

- TCP connect time deltas ($tconn_{CDN} - tconn_{cache}$), probe medians per day, split by address family
- Cache benefits larger over IPv4 than over IPv6



- Longitudinally: benefits of caches grow over IPv4, however, cache benefits become smaller over IPv6
- Latency similar when caches dual-stacked; however, benefits more pronounced for IPv4-only caches compared to IPv6-only caches

Latency benefits toward ISP caches have become more pronounced over IPv4 and less pronounced over IPv6 through the years.

Conclusions

Conclusions: Limitations

- Caches identified by matching ASNs, does not account for peering
- Location of probes biased toward EU and NA
  - Limits analysis of regional differences

Conclusions: Summary

- Latency and IP path lengths similar between both address families
  $\Rightarrow$ High IPv6 preference, however, slight drops during peak hours
- Throughput increases over the years, though not specific to Netflix
  - Over IPv6 lower than over IPv4 in most cases (>75%)
  - However, most deltas only within ± 1 MB/s
- Caches reachable within 6 IP hops and 20 ms
  - IP path lengths shorter by 40–50%
  - Latency lower by 64%
  - Throughput higher by factor of three
  - Latency benefits more pronounced over IPv4 compared with IPv6

Dataset and analysis scripts online:

`https://github.com/tv-doan/infocom-2020-netflix`

Trinh Viet Doan
`doan@in.tum.de`

Vaibhav Bajpai
`bajpaiv@in.tum.de`

Sam Crawford
`sam@samknows.com`