Understanding the Impact of Network Infrastructure Changes using Large-Scale Measurement Platforms

Vaibhav Bajpai
TU Munich
`bajpaiv@in.tum.de`

Jürgen Schönwälder
Jacobs University Bremen
`j.schoenwaelder@jacobs-university.de`

***Abstract*—A number of large-scale network measurement platforms have emerged in the last few years. These platforms have deployed thousands of measurement probes at strategic locations within the access and backbone networks and at residential gateways. The primary goal of these efforts is typically to measure the performance of broadband access networks and to help regulators sketch better policy decisions. In this dissertation we expand the goal further by using large-scale measurement platforms to understand the impact of network infrastructure changes. We deploy probes at the edge of the network to measure IPv6 performance and to dissect last-mile latency characteristics of access networks.**

I. INTRODUCTION

A *large-scale measurement platform* is an infrastructure of dedicated hardware probes that periodically run network measurements tests. These platforms have been deployed to satisfy specific use-case requirements. For instance, a number of platforms (such as CAIDA Archipelago (Ark) [14] and DIMES [15]) emerged in the past to accurately map the network topology of the Internet. Several years of research efforts has matured this area. Recently we have seen a shift towards deployment of performance measurement platforms that measure fixed-line (such as SamKnows and BISmark [16]) networks performance and provide network operational support (such as RIPE Atlas [17] and PerfSONAR). This has been motivated [18] by the emerging need to not only assess the broadband quality but also to verify service offers against contractual agreements. Sundaresan *et al.* in [19] (2011) have used measurement data from such performance measurement platforms (a swarm of deployed SamKnows probes) to investigate the throughput and latency of access network links across multiple ISPs in the United States. They have analyzed this data together with data from the BISmark platform [16] to investigate different traffic shaping policies enforced by ISPs and to understand the bufferbloat [20] phenomenon. The empirical findings of this study have been repraised by Canadi *et al.* in [21] (2012) where they use crowdsourced data from `speedtest.net` to compare both results.

The primary aim of all these activities is to measure the performance and reliability of broadband access networks and facilitate the regulators with research findings to help them make policy decisions. In this dissertation [1], we expand the goal by using large-scale measurement platforms to *understand the impact of network infrastructure changes*.

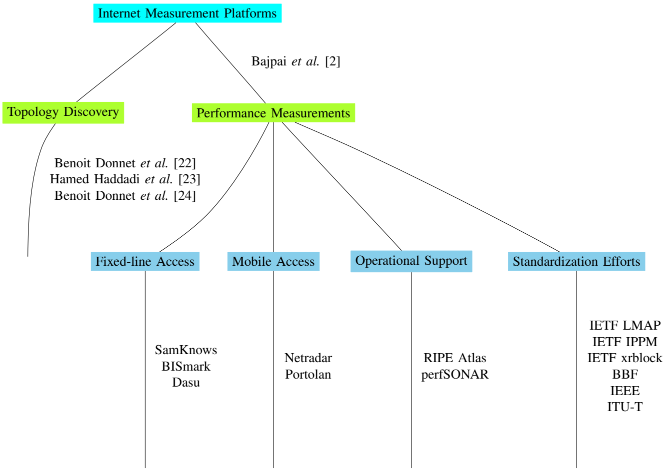

Fig. 1. *An outline of the survey on Internet performance measurement platforms and related standardization efforts.*

We deploy probes at the edge of network to measure IPv6 performance and to dissect last-mile latency characteristics of access networks. This dissertation [1] largely provides three main contributions –

- **Survey on performance measurement platforms** – Initially, measurement platforms were deployed to measure the topology of the Internet. Such topology measurement platforms have been surveyed in the past [22], [23], [24]. In the last couple of years, this focus has evolved towards the measurement of network performance. This has been supported by the deployment of a number of performance measurement platforms. We provide a survey (see § II) of such Internet performance measurement platforms [2]. For each performance measurement platform (see Fig. 1), we present its coverage, scale, lifetime, deployed metrics and measurement tools, architecture and overall research impact. Furthermore, we discuss standardization efforts that are currently being pursued in this space.
- **Measuring IPv6 performance** – A large focus of IPv6 measurement studies in the past has been on measuring IPv6 adoption [25], [26], [27] on the Internet. However, there has been very little to no study [28] on measuring IPv6 performance. We measure IPv6 performance from

the edge of the network (see § III) to popular content delivery services on the Internet. We present metrics, measurement tools, measurement insights and experience from studying geographically varied IPv6 networks. We provide a comparison of how content delivery [3], [4], [5] over IPv6 compares to that of IPv4. We also identify and document glitches [6] in this content delivery that can help improve user experience over IPv6. Our longitudinal observations also identify areas of improvements [7], [8], [9] in the standards work for the IPv6 operations community at the Internet Engineering Task Force (IETF).

- **Dissecting last-mile latency characteristics** – Last-mile latency is a key broadband network performance indicator. However little is known [19], [44] about the characteristics of last-mile latency in access networks. We perform a characterization [1] of last-mile latency (see § IV) by time of day, by subscriber location, by broadband product subscription and by access technology used by the DSL modem. We show that DSL deployments not only tend to enable interleaving on the last-mile, but also employ multiple depth levels that change over time. Our characterization of last-mile latency can be used by simulation studies to model DSL, cable and fibre access links in the future.

II. INTERNET PERFORMANCE MEASUREMENT PLATFORMS

Recently we have seen a trend towards the deployment of Internet performance measurement platforms that provide network operational support and measure fixed-line and mobile access networks. This has been motivated by the emerging need to not only assess the broadband quality but also to verify service offers against contractual agreements. Platforms focussing on inferring the Internet topology have been surveyed in the past [22], [23], [24]. Metrics and tools usually employed in active measurements have also been surveyed [29], [30]. However, there has been no survey on Internet performance measurement platforms.

In [2] (2015), we present a taxonomy of Internet measurement platforms. We subdivide them into topology discovery and performance measurement platforms and further classify the performance measurement platforms based on their deployment use-case – fixed-line access measurements, mobile access measurements and operational support as shown in Fig. 1. We describe performance measurement platforms in detail by exploring their scale, coverage, timeline, deployed metrics and measurement tools, architecture and overall research impact. We also present common set of measurement tools shared by these performance measurement platforms along with the level of collaboration amongst them through the usage of publicly available datasets. We also show how platforms have been using measurement facilitators to conglomerate data from multiple sources to pursue a particular research question. We conclude the survey by describing recent standardization efforts to make large-scale performance measurement platforms interoperable.

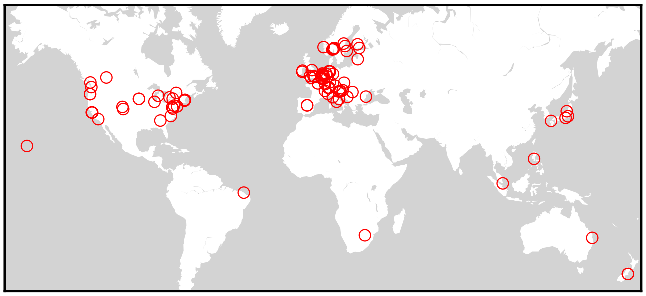

Fig. 2. *Geographical distribution of ~100 dual-stacked SamKnows probes as of Jan 2017. Majority of these probes are deployed in residential settings.*

III. MEASURING IPv6 PERFORMANCE

A large focus of IPv6 measurement studies in the past has been on measuring IPv6 adoption [25], [26], [27] on the Internet. This involved measuring addressing, naming, routing and reachability aspects of IPv6. However, there has been very little work [28] on measuring the performance of delivered services over IPv6. This has largely been due to lack of the availability of content over IPv6. This changed significantly during the span (2013 - 2016) of this dissertation work [1] as a cascading effect of a number of events. For one, the World IPv6 Launch day in 2012 [31] gathered several notable content providers to start providing services over both IPv4 and IPv6. This was also driven by the rapidly exhausting pool of IPv4 address space. As of today, 4/5 RIRs – APNIC (in Apr 2011), RIPE (in Sep 2012), LACNIC (in Jun 2014), and ARIN (in Sep 2015) have exhausted their IPv4 address pool [32] and consequently LIRs now receive allocations from within the last available IPv4 /8 address block. As a result of this depletion, within a span of 3 years, a number of large IPv6 broadband rollouts have also happened [3]. These efforts have eventually led to an increased global adoption of IPv6. For instance, IPv6 adoption jumped during the span of this dissertation work [1] from ~0.85% (as of Sep 2012) to ~11.48% (as of May 2016) according to Google's IPv6 adoption statistics [33]. These numbers demonstrate that IPv6 is no longer an optional IP stack protocol. However, there has been very little to no study [28] on measuring IPv6 performance. This dissertation [1] fills the gap to measure IPv6 performance of operational dual-stacked content services.

We investigated (see § II) potential performance measurement platforms that we could leverage for measuring IPv6 performance. For instance, RIPE Atlas [2] with ~9.1K (~2.4K back in Jan 2013) connected probes with ~2.2K dual-stacked probes [13] as of Jan 2017 is ideal, but is limited in the number of metrics it can measure (primarily `ping` and `traceroute`). PlanetLab vantage points are restricted to mostly research networks and IPv6 support was only recently added to PlanetLab. Therefore, we deployed ~100 SamKnows probes (see Fig. 2) at locations with native IPv6 connectivity A majority of these probes are deployed in residential settings within the RIPE and ARIN region. To put numbers into

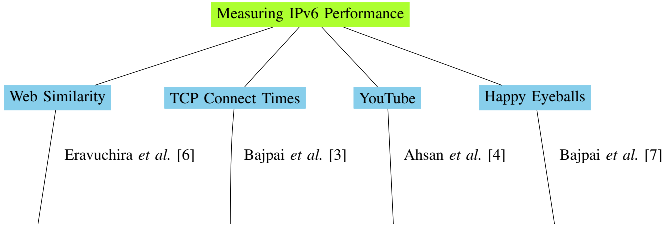

Fig. 3. *An outline on measuring IPv6 performance.*

perspective, this is more than the number of CAIDA Ark probes (75 as of Jan 2017) with native IPv6 connectivity. We use these SamKnows probes to measure IPv6 performance from 66 different origin ASes. Fig. 3 shows an outline of this measurement research.

A. *Web Similarity*

Content providers need to ensure that the content delivered over IPv4 and IPv6 is identical. This is a 2−step process, whereby the content provider has to begin by providing an AAAA record of the service endpoint (or the upfront load balancer) to the DNS resolvers. The end-host then must be able to receive the same content when requesting services from the resolved IPv6 endpoint. IPv6 adoption studies have mostly focussed on the first step by measuring the amount of AAAA entries in DNS resolvers. The similarity of the content served over IPv4 and IPv6 has not been measured in practice. As a result, it remains unclear whether webpages accessed over IPv6 appear similar to their IPv4 counterparts.

In order to address this question, we develop and deploy an active test (`simweb`) that uses well-known content and service complexity metrics [34] to quantify the level of webpage similarity. In [6] (2016), we use this test to measure the similarity of dual-stacked webpages from SamKnows (see Fig. 2) probes. We witnessed that 14% of the ALEXA top 100 dual-stacked websites exhibit dissimilarity in the *number* of fetched webpage elements with 6% showing more than 50% difference. 94% of dual-stacked websites exhibit dissimilarity in *size* with 8% showing at least 50% difference. As such, the content is dissimilar over both address families, but it remains unclear on what factors contribute to this dissimilarity.

In [6], we went further to perform a causal analysis. We found that webpages have higher failure rates over IPv6. For instance, 27% of dual-stacked websites have some fraction of webpage elements that fail over IPv6 with 9% of the websites having more than 50% webpage elements that fail over IPv6. Worse, 6% announce AAAA entries in the DNS but no content is delivered over IPv6 when an HTTP request is made. For instance, `www.bing.com` (a participant of the World IPv6 Launch Day in 2012) [3] is one of the websites that exhibits complete failure over IPv6 because it stopped providing services over IPv6 since Sep 2013. In [6], we show that these failure rates are largely affected by DNS resolution errors on images, javascript and CSS content delivered from both same−origin and cross−origin sources.

B. *TCP Connect Times*

In situations where a webpage element can be fetched over both address families, the performance aspects of retrieving the content over IPv4 and IPv6 need an investigation. This is particularly important because the default address selection policy [35] makes clients prefer connections made over IPv6. However, it's unclear whether users experience benefit (or an added penalty) when connecting to websites over IPv6.

In order to address this question, we introduce a metric and an implementation (`happy`) that measures TCP connection establishment times. By repeated execution of `happy`, we are able to collect time series of TCP connect times that provide us with insights on how IPv6 connectivity to websites compares to that of IPv4. In [3] (2015), we use this test to measure TCP connect times to 100 dual-stacked websites from SamKnows (see Fig. 2) probes. Observations of TCP connect times using an year-long dataset (2013−2014) revealed that latency over IPv6 (back in 2013) was considerably worse. In order to identify the reason for this degraded performance, we went further to cluster websites by CDN deployments and observed that these CDN clusters were different for IPv4 and IPv6. This revealed that the disparity in latency was because popular websites were served over IPv4 from CDN caches deployed directly within access networks, but such caches were largely absent for IPv6. This lead to relatively higher TCP connection establishment times over IPv6. We revisited this question in [7] (2016) and witnessed that TCP connect times to popular websites have considerably improved over the last 3 years (2013−2016). In fact, as of May 2016, 18% of ALEXA top 10K websites are now faster over IPv6 with 91% of the rest being at most 1 ms slower. We found that in such a changed landscape, Google now employs prefix blacklists [3] to block hosts behind resolvers from receiving their services over IPv6 in situations where latency over IPv6 is considerably (100 ms or more) worse than IPv4.

C. *YouTube*

We went further to study the IPv6 performance of a specific workload on the Internet. Studies have shown that IPv6 traffic is largely dominated by services running over HTTP and YouTube is the primary service over HTTP that contributes heavily to large volumes of IPv6 traffic. However, it's not known how often does streaming a YouTube video over IPv6 fail and how does this failure rate compare to that of IPv4.

In order to address this question, we deploy an active test (`youtube`) on the SamKnows (see Fig. 2) probes. Using a 2-year (2014−2016) long dataset we showed that success rates of streaming a stall-free version of the video over IPv6 were lower compared to that of IPv4 but they tend to have improved over time. Although, in situations where streaming succeeds over both address families, it remains unknown whether users experience benefit (or an added penalty) when streaming YouTube videos over IPv6. In [4], we observed consistently higher TCP connect times and startup delays (~100 ms or more) over IPv6. Furthermore, throughput achieved was also consistently lower over IPv6 for both audio and video streams.

Although we witnessed low stall rates over both address families and reduced stall durations over the years, in situations where a stall occurred, the stall durations were relatively higher (1s or more) over IPv6. This raises questions on what factors contribute towards the worse streaming performance over IPv6. In [4], we showed that the performance difference is due to disparity in the availability of content caches, whereby content caches over IPv6 are largely absent.

D. Happy Eyeballs

Using our longitudinal observations, we try to identify areas of improvements in the standards work for the IPv6 operations community at the IETF. The Happy Eyeballs (HE) algorithm [36] (2012) for instance, provides recommendations to application developers to help prevent bad user experience in situations where IPv6 connectivity is broken. The algorithm when combined with the default address selection policy [35] (2012), tends to give a noticeable advantage (300 ms) to connections made over IPv6. The HE timer value was chosen during a time (2012) when broken IPv6 connectivity was quite prevalent, which made applications stall for several seconds before attempting a connection over IPv4. The broken IPv6 connectivity has been largely attributed to failures caused by Teredo [37] and 6to4 relays [38]. However, Teredo/6to4 technologies have seen a rapid decline over the years (~0.01% as of Jan 2017) due to efforts made by the IPv6 operations community. In such a changed landscape, the effects of the HE timer value (300 ms) on the overall experience of a dual-stacked user remains largely unclear. For instance, it's unclear how often HE makes a bad decision of choosing IPv6 when it is slower and in such situations what is the amount of imposition (in terms of latency impact) a dual-stacked user has to pay as a result of the high HE timer value.

In order to address this question, we measured the effects of the HE algorithm. In [9] (2013), we showed that HE (with a 300 ms timer value) never prefers IPv6 using Teredo except in situations where IPv4 reachability of the destination endpoint is broken. We went further in [7] (2016) and showed that only ~1% of the TCP connect times over IPv6 (2013 - 2016) were ever above the HE timer value (300 ms), which leaves ~2% chance for IPv4 to win a HE race towards these websites. As such, IPv6 connections to 99% of ALEXA top 10K websites were preferred more than 98% of the time, although in ~90% of the cases, slower IPv6 connections were preferred by HE. In [4], we also witnessed that HE strongly prefers (more than 97%) connections made over IPv6 for streaming YouTube even though this preference to IPv6 brings worse performance in comparison with IPv4. As such, HE timer value has passed its time, but it remains unclear on what shall be the new HE timer value that can provide the same preference levels over IPv6 as is today but also reduces the performance penalty in situations where IPv6 is considerably slower. In order to address this question, in [7], we went a step further and showed that reducing the HE timer value from 300 ms to 150 ms provides a margin benefit of 10% (in terms of latency) while retaining same IPv6 preference levels for 99% of the measured dual-stacked websites.

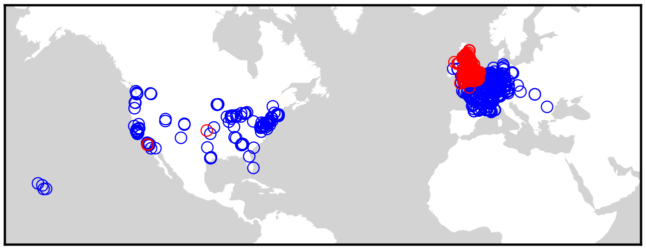

Fig. 4. *Distribution of 696 RIPE Atlas v3 (blue) and 1245 SamKnows (red) residential probes. RIPE Atlas probes span the EU (521) and the US (161), while SamKnows probes span the UK (1233) and the US (11).*

IV. DISSECTING LAST-MILE LATENCY CHARACTERISTICS

Measurement studies [39] (2013) performed using the BISmark [2], [16] platform, have shown that latency becomes a critical factor impacting QoE in networks where downstream throughput exceeds 16 Mb/s. This has driven content providers to deploy content caches in service provider networks to move the content as close [40] to the edge as possible. Furthermore, recent [41] (2015) and upcoming standards [42], [43] (2017) cater to this requirement to target operation at a much reduced latency. It was recently shown [39] (2013) that last-mile latency is a major contributor to end-to-end latency and it contributes heavily to DNS lookup and page load times. Last-mile latency is becoming a key broadband network performance indicator. However little is known [19], [44] (2011, 2007) about the characteristics of last-mile latency in access networks. In order to develop a methodology for last-mile latency measurements, we ask whether last-mile latency should include latencies within the home network.

In [1] we ran month-long `traceroute` measurements using residential 696 RIPE Atlas [2] and 1245 SamKnows [2] probes as shown in Fig. 4. We witnessed 19.2% of RIPE Atlas probes and 29.7% of SamKnows probes exhibit $hop1$ latency contributing to 10% or more of $hop2$ latency. As such, the home network latency can make a discernible contribution and therefore should not be accounted when measuring last-mile latency. We went further and asked whether queuing delay caused by bufferbloat on home routers has an impact on last-mile latency. In [1], we found that 9.95% of SamKnows probes show $hop1/hop2$ contribution of more than 100% where $hop1$ latencies for these probes appear considerably stable at ~50 ms. These probes are connected to home routers that rate limit ICMP responses to TTL expiry and therefore vantage points connected to these home routers should not be used for baseline last-mile latency measurements.

Using this methodology we further study the characteristics of last-mile. For instance, it's known [19] that DSL networks enable interleaving on the last-mile which increases last-mile latencies for DSL users. However, it's unclear whether ISPs employ multiple interleaving depth levels and if these depth levels vary with time. In [1] we show that some DSL providers dynamically adapt interleaving depth levels depending on the line characteristics and geographic location of the subscriber. For some measurement points, we observed depth level changes occuring on a weekly time scale. Using this knowledge, we ask further on what factors does last-mile latency depend and how it varies over the time of the day.

In [1] we show that once the effects of queuing delay caused by bufferbloat have been eliminated, access networks tend to exhibit robust last-mile latency. We witnessed that last-mile latencies of a service provider can depend on the geographic location of a subscriber. We observed significant last-mile latency differences for US cable service providers across the east (centered at ~32ms) and west (centered at ~8ms) coast. We showed that last-mile latencies of DSL deployments vary with the the broadband product subscription, whereby last-mile latencies for products based on ADSL2+ and VDSL are significantly lower compared to the latency of ADSL1 products. Finally, we ask whether it would be possible to recommend a characteristic value of last-mile latency by access technology.

In [1], we witnessed that last-mile latency for DSL deployments is centered ~16 ms. Cable networks show a last-mile latency centered ~8 ms and fibre to the home networks show a last-mile latency centered ~4 ms. These characteristics values can now be used by simulation studies to appropriately model DSL, cable and fibre access links.

V. RELEVANCE AND IMPACT

The survey [2] on performance measurement platforms (see § II) is relevant for parties who build and maintain large-scale measurement platforms. The survey is also useful to early researchers to get acquainted with the background in measurement-based research. Parties involved in large-scale measurement standardization activities [10], [11] may also find this contribution useful.

The study on measuring IPv6 performance (see § III) is relevant for network operators that are either in the process of or are in early stages of IPv6 deployment. This research provides content providers insights towards how their service delivery compares over IPv6 to that of IPv4.

The study on dissecting last-mile latency characteristics (see § IV) extends our understanding of last-mile latency witnessed by home users. CDN providers that attempt to optimise content delivery towards the edge of the network may benefit from the identified characteristics of the last-mile. This work will also benefit service providers since it promotes the possibility of caching popular content near to the home routers to further eliminate the bottlenecks induced by last-mile latency. This research may also serve as possible input for ongoing standardization efforts such as QUIC [42] and TLS 1.3 [43] within the IETF that attempt to target operations at much reduced latency.

Lessons learned [12] from pursuing this dissertation may prove valuable to the wider measurement community in general. As a side contribution, our vantage point selection methodology [13] to identify home probes in the RIPE Atlas platform can serve as a good starting point for future broadband measurement studies using the RIPE Atlas platform.

VI. FUTURE DIRECTIONS

In [3] we measured TCP connect times to dual-stacked websites. It would be nice to know how does the raw throughput performance of a TCP connection towards a dual-stacked website compare over IPv6 to that of IPv4. This requires measuring the Bulk Transfer Capacity (BTC) (since we observe TCP). This is not straightforward since BTC measurement tools require access at both ends of the measured path. As such, this effort will require collaborative support from large CDN providers. In [3] we also showed that TCP connect times over IPv6 to popular dual-stacked websites have considerably improved over time. However, it is unclear whether this is due to IPv6 content moving closer to the client (similar to how it is in IPv4). Moreover, in situations where there is considerable disparity in TCP connect times to the same website, it remains unclear whether this is due to dissimilarity of paths traversed over IPv4 and IPv6. It would be nice to measure the similarity of paths traversed for each website. In [7] we showed that lowering the HE timer value to 150 ms (from 300 ms) provides a margin benefit of 10% while retaining similar IPv6 preference levels. Another approach is to make clients adaptively change the HE timer value based on the previously witnessed history of the TCP connect times over both address families. However, it remains unclear whether browser implementations prefer to trade complexity for such an increased intelligence.

We also showed that last mile latencies of a service provider can depend on the geographic location of a subscriber. We observed significant last-mile latency differences for US cable service providers across the east and west coast. However, the causes of this observed effect remain unclear. Analyzing this further requires collaboration with service providers.

PUBLICATIONS

[1] V. Bajpai, "Understanding the Impact of Network Infrastructure Changes using Large-Scale Measurement Platforms," Ph.D. dissertation, Jacobs University Bremen, Germany, 2016. [Online]. Available: http://nbn-resolving.de/urn:nbn:de:gbv:579-opus-1005861

[2] V. Bajpai and J. Schönwälder, "A Survey on Internet Performance Measurement Platforms and Related Standardization Efforts," ser. IEEE Communications Surveys and Tutorials (COMST), vol. 17, no. 3, 2015. [Online]. Available: http://dx.doi.org/10.1109/COMST.2015.2418435

[3] ——, "IPv4 versus IPv6 - who connects faster?" ser. IFIP Networking Conference, Networking 2015, 2015, pp. 1–9. [Online]. Available: http://dx.doi.org/10.1109/IFIPNetworking.2015.7145323

[4] S. Ahsan, V. Bajpai, J. Ott, and J. Schönwälder, "Measuring YouTube from Dual-Stacked Hosts," ser. Passive and Active Measurement (PAM) Conference, 2015, pp. 249–261. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-15509-8_19

[5] V. Bajpai and J. Schönwälder, "Measuring TCP Connection Establishment Times of Dual-Stacked Web Services," ser. Conference on Network and Service Management (CNSM), 2013, pp. 130–133. [Online]. Available: http://dx.doi.org/10.1109/CNSM.2013.6727822

[6] S. J. Eravuchira, V. Bajpai, J. Schönwälder, and S. Crawford, "Measuring Web Similarity from Dual-Stacked Hosts," ser. Conference on Network and Service Management (CNSM), 2016, pp. 181–187. [Online]. Available: http://dx.doi.org/10.1109/CNSM.2016.7818415

[7] V. Bajpai and J. Schönwälder, “Measuring the Effects of Happy Eyeballs,” ser. Applied Networking Research Workshop (ANRW), 2016, pp. 38–44. [Online]. Available: http://doi.acm.org/10.1145/2959424.2959429

[8] V. Bajpai and J. Schönwälder, “Understanding the Impact of Network Infrastructure Changes Using Large-Scale Measurement Platforms,” ser. Conference on Autonomous Infrastructure, Management, and Security (AIMS), 2013, pp. 41–44. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-38998-6_5

[9] V. Bajpai and J. Schönwälder, “Measuring the Effects of Happy Eyeballs,” Internet-Draft draft-bajpai-happy-01, Jul. 2013. [Online]. Available: http://tools.ietf.org/html/draft-bajpai-happy-01

[10] J. Schönwälder and V. Bajpai, “A YANG Data Model for LMAP Measurement Agents,” Internet Engineering Task Force, Internet-Draft draft-ietf-lmap-yang-10, Jan. 2017, work in Progress. [Online]. Available: http://tools.ietf.org/html/draft-ietf-lmap-yang-10

[11] ——, “Using RESTCONF with LMAP Measurement Agents,” Internet Engineering Task Force, Internet-Draft draft-ietf-lmap-restconf-03, Jul. 2016, work in Progress. [Online]. Available: http://tools.ietf.org/html/draft-ietf-lmap-restconf-03

[12] V. Bajpai, S. J. Eravuchira, and J. Schönwälder, “Lessons Learned From Using the RIPE Atlas Platform for Measurement Research,” ser. Computer Communication Review (CCR), vol. 45, no. 3, 2015, pp. 35–42. [Online]. Available: http://doi.acm.org/10.1145/2805789.2805796

[13] V. Bajpai, S. J. Eravuchira, J. Schönwälder, R. Kisteleki, and E. Aben, “Vantage Point Selection for IPv6 Measurements: Benefits and Limitations of RIPE Atlas Tags,” ser. International Symposium on Integrated Network Management (IM), 2017 (to appear).

ACKNOWLEDGEMENTS

This work was funded by Flamingo, a Network of Excellence project (ICT-318488) supported by the European Commission under its Seventh Framework Programme and by Leone, the European Communitys Seventh Framework Programme in (FP9/2007-2013) Grant No. 317647.

REFERENCES

[14] kc claffy, “The 7th Workshop on Active Internet Measurements (AIMS7) Report,” ser. Computer Communication Review (CCR), vol. 46, no. 1, 2016. [Online]. Available: http://doi.acm.org/10.1145/2875951.2875960

[15] Y. Shavitt and E. Shir, “DIMES: Let the Internet Measure Itself,” ser. Computer Communication Review (CCR), vol. 35, no. 5, 2005, pp. 71–74. [Online]. Available: http://doi.acm.org/10.1145/1096536.1096546

[16] S. Sundaresan, S. Burnett, N. Feamster, and W. de Donato, “BISmark: A Testbed for Deploying Measurements and Applications in Broadband Access Networks,” ser. USENIX Annual Technical Conference (ATC), 2014, pp. 383–394. [Online]. Available: https://www.usenix.org/conference/atc14/technical-sessions/presentation/sundaresan

[17] “RIPE Atlas: A Global Internet Measurement Network,” ser. Internet Protocol Journal (IPJ), Sep. 2015. [Online]. Available: http://ipj.dreamhosters.com/wp-content/uploads/2015/10/ipj18.3.pdf

[18] M. Linsner, P. Eardley, T. Burbridge, and F. Sorensen, “Large-Scale Broadband Measurement Use Cases,” RFC 7536 (Informational), Internet Engineering Task Force, May 2015. [Online]. Available: http://www.ietf.org/rfc/rfc7536.txt

[19] S. Sundaresan, W. de Donato, N. Feamster, R. Teixeira, S. Crawford, and A. Pescapè, “Broadband Internet Performance: A View From the Gateway,” ser. ACM SIGCOMM 2011 Conference, 2011, pp. 134–145. [Online]. Available: http://doi.acm.org/10.1145/2018436.2018452

[20] J. Gettys and K. M. Nichols, “Bufferbloat: Dark Buffers in the Internet,” ser. Communications of the ACM (CACM), vol. 55, no. 1, 2012, pp. 57–65. [Online]. Available: http://doi.acm.org/10.1145/2063176.2063196

[21] I. Canadi, P. Barford, and J. Sommers, “Revisiting broadband performance,” ser. Internet Measurement Conference (IMC), 2012. [Online]. Available: http://doi.acm.org/10.1145/2398776.2398805

[22] B. Donnet, “Internet Topology Discovery,” ser. Data Traffic Monitoring and Analysis, 2013, pp. 44–81. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-36784-7_3

[23] H. Haddadi, M. Rio, G. Iannaccone, A. W. Moore, and R. Mortier, “Network topologies: Inference, modeling, and generation,” ser. IEEE Communications Surveys and Tutorials (COMST), vol. 10, 2008. [Online]. Available: http://dx.doi.org/10.1109/COMST.2008.4564479

[24] B. Donnet and T. Friedman, “Internet topology discovery: A survey,” ser. IEEE Communications Surveys and Tutorials (COMST), 2007. [Online]. Available: http://dx.doi.org/10.1109/COMST.2007.4444750

[25] L. Colitti, S. H. Gunderson, E. Kline, and T. Refice, “Evaluating IPv6 Adoption in the Internet,” ser. Passive and Active Measurement Conference (PAM), 2010, pp. 141–150. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-12334-4_15

[26] A. Dhamdhere, M. J. Luckie, B. Huffaker, kc claffy, A. Elmokashfi, and E. Aben, “Measuring the deployment of IPv6: topology, routing and performance,” ser. Internet Measurement Conference (IMC), 2012. [Online]. Available: http://doi.acm.org/10.1145/2398776.2398832

[27] J. Czyz, M. Allman, J. Zhang, S. Iekel-Johnson, E. Osterweil, and M. Bailey, “Measuring IPv6 adoption,” ser. SIGCOMM, 2014, pp. 87–98. [Online]. Available: http://doi.acm.org/10.1145/2619239.2626295

[28] M. Nikkhah, R. Guérin, Y. Lee, and R. Woundy, “Assessing IPv6 rough Web Access - A Measurement Study and Its Findings,” ser. Conference on Emerging Networking Experiments and Technologies (CoNEXT), 2011. [Online]. Available: http://doi.acm.org/10.1145/2079296.2079322

[29] F. Michaut and F. Lepage, “Application-oriented network metrology: Metrics and active measurement tools,” ser. IEEE Communications Surveys and Tutorials (COMST), vol. 7, no. 1-4, 2005, pp. 2–24. [Online]. Available: http://dx.doi.org/10.1109/COMST.2005.1610543

[30] R. Prasad, C. Dovrolis, M. Murray, and K. Claffy, “Bandwidth Estimation: Metrics, Measurement Techniques, and Tools,” ser. IEEE Network, vol. 17, no. 6, Nov 2003, pp. 27–35. [Online]. Available: http://dx.doi.org/10.1109/MNET.2003.1248658

[31] The Internet Society, “World IPv6 Launch,” http://www.worldipv6launch.org, [Online; accessed 11-January-2016].

[32] P. Richter, M. Allman, R. Bush, and V. Paxson, “A Primer on IPv4 Scarcity,” ser. Computer Communication Review (CCR), vol. 45, no. 2, 2015. [Online]. Available: http://doi.acm.org/10.1145/2766330.2766335

[33] “Google IPv6 Adoption Statistics,” http://www.google.com/intl/en/ipv6/statistics.html, [Online; accessed 11-January-2016].

[34] M. Butkiewicz, H. V. Madhyastha, and V. Sekar, “Characterizing Web Page Complexity and Its Impact,” *IEEE/ACM Trans. Netw.*, vol. 22, no. 3, pp. 943–956, Jun. 2014. [Online]. Available: http://dx.doi.org/10.1109/TNET.2013.2269999

[35] D. Thaler, R. Draves, A. Matsumoto, and T. Chown, “Default Address Selection for Internet Protocol Version 6 (IPv6),” RFC 6724, Sep. 2012. [Online]. Available: http://www.ietf.org/rfc/rfc6724.txt

[36] D. Wing and A. Yourtchenko, “Happy Eyeballs: Success with Dual-Stack Hosts,” RFC 6555 (Proposed Standard), Internet Engineering Task Force, Apr. 2012. [Online]. Available: http://www.ietf.org/rfc/rfc6555.txt

[37] C. Huitema, “Teredo: Tunneling IPv6 over UDP through Network Address Translations (NATs),” RFC 4380, Internet Engineering Task Force, Feb. 2006. [Online]. Available: http://www.ietf.org/rfc/rfc4380.txt

[38] B. Carpenter and K. Moore, “Connection of IPv6 Domains via IPv4 Clouds,” RFC 3056 (Proposed Standard), Internet Engineering Task Force, Feb. 2001. [Online]. Available: http://www.ietf.org/rfc/rfc3056.txt

[39] S. Sundaresan, N. Feamster, R. Teixeira, and N. Magharei, “Measuring and mitigating web performance bottlenecks in broadband access networks,” ser. Internet Measurement Conference (IMC), 2013, pp. 213–226. [Online]. Available: http://doi.acm.org/10.1145/2504730.2504741

[40] X. Fan, E. Katz-Bassett, and J. S. Heidemann, “Assessing Affinity Between Users and CDN Sites,” ser. Traffic Monitoring and Analysis, 2015. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-17172-2_7

[41] M. Belshe, R. Peon, and M. Thomson, “Hypertext Transfer Protocol Version 2 (HTTP/2),” RFC 7540, Internet Engineering Task Force, May 2015. [Online]. Available: http://www.ietf.org/rfc/rfc7540.txt

[42] J. Iyengar and M. Thomson, “QUIC: A UDP-Based Multiplexed and Secure Transport,” Internet Engineering Task Force, Internet-Draft draft-ietf-quic-transport-01, Jan. 2017, work in Progress. [Online]. Available: http://tools.ietf.org/html/draft-ietf-quic-transport-01

[43] E. Rescorla, “The Transport Layer Security (TLS) Protocol Version 1.3,” Internet Engineering Task Force, Internet-Draft, Oct. 2016. [Online]. Available: http://tools.ietf.org/html/draft-ietf-tls-tls13-18

[44] M. Dischinger, A. Haeberlen, P. K. Gummadi, and S. Saroiu, “Characterizing residential broadband networks,” ser. Internet Measurement Conference (IMC), 2007, pp. 43–56. [Online]. Available: http://doi.acm.org/10.1145/1298306.1298313