Measuring IPv6 Adoption

Jakub Czyz
University of Michigan

Mark Allman
International Computer Science Institute

Jing Zhang
University of Michigan

Scott Iekel-Johnson
Arbor Networks, Inc.

Eric Osterweil
Verisign Labs

Michael Bailey
University of Michigan

ABSTRACT

After several IPv4 address exhaustion milestones in the last three years, it is becoming apparent that the world is running out of IPv4 addresses, and the adoption of the next generation Internet protocol, IPv6, though nascent, is accelerating. In order to better understand this unique and disruptive transition, we explore twelve metrics using ten global-scale datasets to create the longest and broadest measurement of IPv6 adoption to date. Using this perspective, we find that adoption, relative to IPv4, varies by two orders of magnitude depending on the measure examined and that care must be taken when evaluating adoption metrics in isolation. Further, we find that regional adoption is not uniform. Finally, and perhaps most surprisingly, we find that over the last three years, the nature of IPv6 utilization—in terms of traffic, content, reliance on transition technology, and performance—has shifted dramatically from prior findings, indicating a maturing of the protocol into production mode. We believe IPv6's recent growth and this changing utilization signal a true quantum leap.

Categories and Subject Descriptors

C.2.5 [**Local and Wide-Area Networks**]: Internet

Keywords

Internet; IP; IPv4; IPv6; DNS; Measurement

1. INTRODUCTION

IPv4 is the common thread that has held the Internet together since its very early years, and, thus, it is both the most important and most widely-deployed networking protocol in existence. For basic end-to-end connectivity, devices need to have a unique IP address, but the world is rapidly running out of available IPv4 address space. Thus, if we want the Internet to continue growing and delivering its societal and economic benefits for the next generation, we have a challenge. In just the three years since February 2011, the organizations responsible for allocating IPv4 addresses, including the Internet Assigned Numbers Authority (IANA), as well as two of its five subordinate regional Internet registries (RIRs) have either completely exhausted address space or resorted to rationing their final address block. It appears that after years of both architecture changes and stop-gap measures forestalling IPv4 address exhaustion (e.g., classless interdomain routing [CIDR] [12], network address translation [NAT] [37]) the Internet has now begun its first core protocol change. Exhaustion is providing increased impetus for the network to finally adopt the next version of IP, IPv6. We show that, while raw IPv6 Internet traffic is still a small fraction (0.64%), the *nature* of its use and the *trajectory* of growth have shifted dramatically, and, consequently, IPv6 should no longer be dismissed by researchers as an uninteresting rarity.

In this paper, our aim is to empirically understand the adoption of IPv6. A once-in-a-lifetime opportunity to observe technological change on such a grand scale, this is both practically and scientifically important. However, with a handful of exceptions, most of the individual assessments of IPv6 that the community has produced, to date, are anecdotal (e.g., using one server's, campus's or Internet exchange point's perspective) and/or focus on only a single aspect of IPv6 adoption (e.g., route advertisement). No single existing perspective or study suffices to truly gauge the state of IPv6 in the large. We argue that, while this previous body of work has been invaluable at tracking aspects of adoption, a broader approach to measuring IPv6 deployment, which assembles a breadth of observations and compares existing datasets against each other, is needed to truly understand where we are. The goal is not to pick the "best" one, but to understand the *systemic* state of IPv6 deployment via all available measurements.

To achieve our goal we assemble a set of six publicly-accessible datasets that speak to one or more aspects of IPv6 adoption. We add four additional, previously-unpublished datasets, including a global Internet traffic dataset that includes traffic statistics from 260 providers and represents 16,200 petabytes/month, or approximately 33-50% of all Internet traffic—the largest traffic sample reported in an IPv6 study. In addition to the traffic data, we add DNS query data from several of the largest globally-distributed IPv4-based replicas of the .com and .net top-level domains (TLDs), as well as nearly all native IPv6 replicas for these TLDs.

To create a comprehensive view of IPv6 adoption, we enumerate and compare both previously-reported and novel metrics and weave them into a taxonomy. In applying these metrics to our datasets, we seek to reveal relationships and patterns not otherwise visible when studying any single metric in isolation. Further, we speculate that several recent IPv4 exhaustion events (IANA, APNIC, RIPE) and community IPv6 "flag days" (World IPv6 Day 2011 and Launch 2012) may have noticeably influenced the progression of adoption.

Through the lens of this comprehensive approach, we compose a picture of the current state of global IPv6 adoption. We find:

SIGCOMM'14, August 17–22, 2014, Chicago, IL, USA.

http://dx.doi.org/10.1145/2619239.2626295.

- **IPv6 is real.** Over the last three years, IPv6 has reached a significant developmental milestone and is finally being used natively and for normal, production traffic, on a non-trivial and accelerating scale. While IPv6 is still under 1% of Internet traffic, it has increased over 400% in each of the last two years, relies much less on transition technologies, and is used for similar applications to IPv4, with similar performance.

- **Measurements vary widely.** IPv6 adoption level differs by up to *two orders of magnitude* depending on the metric used. For instance, although IPv6 monthly address allocations are about 57% of IPv4, the percentage of IPv4 nameservers in .com that are IPv6-reachable is around 3%.

- **Geographic adoption differs.** Global regions are adopting IPv6 at different rates, but the relative level of adoption in the regions also varies by metric, suggesting that incentives and barriers to adopt the new protocol do not just vary globally, but across layers of the network within each region.

2. RELATED WORK

There are many papers in the literature that offer valuable data on the IPv6 adoption process from various perspectives. Several studies characterize IPv6 traffic from the perspective of one or more ISPs (e.g., [23, 34, 36]) and 6to4 relays (e.g., [18, 35]). On June 8, 2011, the Internet Society sponsored "IPv6 World Day" [22] and several pieces of work explore this event explicitly (e.g., [34]). Other work examines IPv6 adoption from the perspective of the World Wide Web (e.g., [7, 29]). Additionally, a variety of contributions explore the technical, economic, and social factors that influence adoption (e.g., [16, 20]). Finally, much previous work focuses on topology measurements and performance in IPv6 and their relationships to IPv4 (e.g., [5, 10, 13, 31, 39, 43, 44]). In contrast to much of these studies, our work sacrifices depth for breadth in order to understand the big picture of IPv6 adoption.

Claffy [6] discusses IPv6 evolution and observes that "we lack not only a comprehensive picture of IPv6 deployment, but also consensus on how to measure its growth, and what to do about it." Our paper is in part a response to this call, offering a possible way forward. Closest to our work in both spirit and substance is Karpilovsky *et al.* [23], who provide a snapshot of IPv6 adoption from three main perspectives (allocation data, routing data, and traffic from a tier-1 ISP). In comparison, our work broadens the traffic perspective to a large sample of global tier-1 ISPs and nearly 100 tier-2/regional ISPs (260 providers in total), includes large samples of .com and .net TLD data, and juxtaposes these datasets with seven additional (mostly public) datasets.

Our DNS packet analysis in § 5 extends work by researchers at Verisign [30, 40]. The key distinction of our contribution here over this previous work is that we examine DNS queries via both IPv4 and IPv6 traffic, we focus on IPv6 adoption, and the data presented is more recent; the earlier Verisign work contains longitudinal IPv4-only traffic analysis, though performed at greater detail.

3. OUR APPROACH

Since our aim is for a comprehensive picture of adoption, we must first decide what aspects to study. We start by thinking about the Internet Protocol from the perspective of the three major types of Internet stakeholders: content providers, service providers, and content consumers. Although there are notable entities that straddle or defy these labels (e.g. vendors and policy makers), these three categories encapsulate the key perspectives we believe should be considered to realistically assess deployment. We next divide the key aspects of IP itself into two classes. The first is the prerequisite functions that IP performs and that must be in place for nodes to communicate, including addressing, naming, routing, and end-to-end reachability. The second class is operational characteristics that are only evident once the network begins forwarding packets, these include transition technology use, traffic volume, application mix, and performance.

In Table 1 we propose one or more *metrics* that characterize the adoption of IPv6 from the key viewpoints sketched above. Some of these metrics cover more than one branch in the taxonomy. We admit that our use of the term "metric" is somewhat loose. Our aim is to point to many aspects of adoption that should be measured, but whose granularity and specificity varies. Thus, each metric could itself be thought of more as a category or issue for which specific measurements should be obtained. In this paper, we present one or several such measurements for each metric that we defined.

While we believe we have identified a sufficiently comprehensive set of metrics to provide a broad picture of adoption, we do not claim completeness. There are countless possible metrics that can tell a coherent and insightful story of the adoption process. Further, while a metric such as performance naturally breaks down into submetrics for assessing delay, loss, jitter, reordering, throughput, etc., the specific facets of IPv6 operation that are important in any given context are likely to vary by application. As such, we do not mean to discourage further assessment along different axes or granularities than we take in this paper. Rather, our goal is to set a course for developing a *high-level* and *comprehensive* understanding of the recent IPv6 adoption process. To this end, we bring to bear several large original datasets, as well as several public or previously published results, and use them to report measurements that align with our taxonomy. Table 2 summarizes the datasets we analyzed, separated into the public ones we reproduce and update, and the unique ones we contribute. We next discuss analyses of these data in detail, showing how adoption level differs and varies as we move from left to right in Table 1 and over time.

4. ADDRESSING

We first examine the initial steps for IPv6 deployment: address allocation and network advertisement.

A1: Address Allocation

Before wide-scale IPv6 communication is possible, IPv6 addresses must be broadly available. Therefore, our first assessment is of the status of IPv6 address allocation. The allocation hierarchy begins with the Internet Assigned Numbers Authority (IANA) allocating address blocks to the five regional Internet registries (RIRs). In turn, the RIRs make allocations to various national and local registries and ISPs. Each RIR publishes a daily snapshot of the blocks of IP (v4 and v6) addresses (i.e., the number of prefixes) allocated to entities below it in the hierarchy. We have captured ten years of these snapshots, starting in January 2004. As a minor caveat, note that the size of a typical IPv6 prefix (2^{96}) is much larger than that of an IPv4 prefix (2^{10}), thus, prefix-based comparisons should be made with caution. However, address allocations typically correspond to network deployments, no matter the protocol; so, relative allocations do shed light on protocol deployment.

Figure 1 shows the aggregate number of prefixes allocated each month across all RIRs. There were less than 30 IPv6 prefixes allocated per month prior to 2007, generally increasing thereafter. In the past several years, we typically find more than 300 prefixes allocated per month, with a high point of 470 prefix allocations in February 2011. By January 2004 there had been 650 IPv6 prefix allocations, while at the end of December 2013 we observe 17,896 to-

Table 1: IPv6 adoption metric taxonomy, showing the main three perspectives that Internet stakeholders occupy, the prerequisites for IPv6 to be used, as well as the operational characteristics of the protocol, once deployed.

		Prerequisite IP Functions				Operational Characteristics	
		Addressing	**Naming**	**Routing**	**End-to-End Reachability**	**Usage Profile**	**Performance**
Perspective	**Content Provider**		N1: Nameservers; R1: Server Readiness		R1: Server Readiness	U3: Transition Technologies	
	Service Provider	A1: Address Allocation; A2: Address Advertisement	N2: Resolvers	A2: Address Advertisement; T1: Topology		U1: Traffic Volume; U3: Transition Technologies	P1: Network RTT
	Content Consumer		N3: Queries		R2: Client Readiness	U2: Application Mix; N3: Queries	

Table 2: Dataset summary showing the time period, scale, and public or new status of the datasets we analyzed.

Dataset	Metrics	Time Period	Recent Scale	Public?
RIR Address Allocations	A1	Jan 2004 – Jan 2014	≈18K allocation snapshots (5 daily)	Yes
Routing: Route Views	A2, T1	Jan 2004 – Jan 2014	45,271 BGP table snapshots	
Routing: RIPE	A2, T1	Jan 2004 – Jan 2014		
Google IPv6 Client Adoption	R2, U3	Sep 2008 – Dec 2013	millions of daily global samples	
Verisign TLD Zone Files	N1	Apr 2007 – Jan 2014	daily snapshots of ≈2.5 million A+AAAA glue records (.com & .net)	
CAIDA Ark Performance Data	P1	Dec 2008 – Dec 2013	≈10 million IPs probed daily	
Arbor Networks ISP Traffic Data	U1, U2, U3	Mar 2010 – Dec 2013	≈33-50% of global Internet traffic; 2013 daily median: 50 terabits/sec (avg.)	No
Verisign TLD Packets: IPv4	N2, N3	Jun 2011 – Dec 2013	4 global sites, 5 of 13 gTLD NS letters (.com/.net), ≈4.5Bn queries	
Verisign TLD Packets: IPv6	N2, N3	Jun 2011 – Dec 2013	15 global sites, both gTLD NS letters (.com/.net) w/IPv6, 647M queries	
Alexa Top Host Probing	R1	Apr 2011 – Dec 2013	10,000 servers probed twice/month	

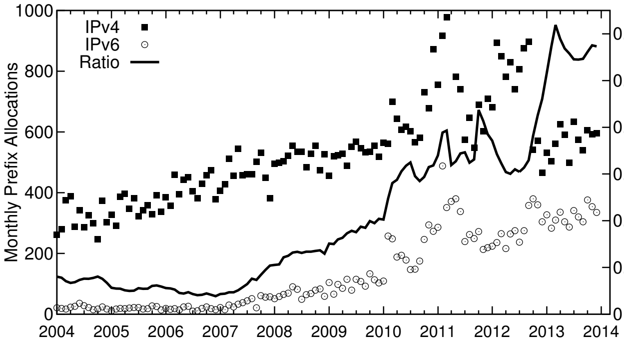

Figure 1: Prefixes allocated. IPv4 and IPv6 allocations accelerate leading up to IANA exhaustion in early 2011. IPv4 dropped in 2012 and was flat in 2013, while IPv6 trended upward.

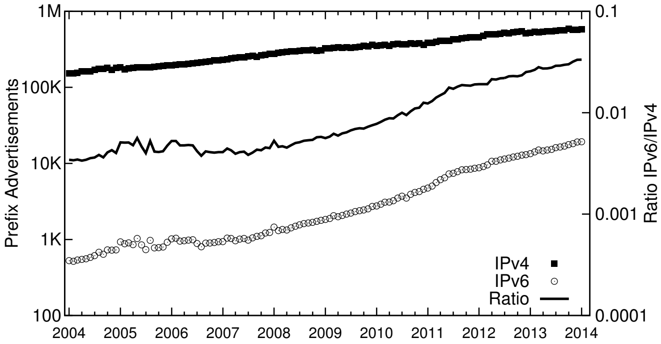

Figure 2: Number of advertised prefixes. Over ten years, IPv6 prefixes increase 37-fold, while IPv4 increase four-fold.

tal prefix allocations—an increase of 27-fold. Finally, we note that at the end of our dataset the allocated IPv6 prefixes cover 2^{113} (i.e., 1.1×10^{34}) addresses.

To put the IPv6 allocation data in context, Figure 1 also shows IPv4 prefix allocations over the same period. The number of IPv4 prefix allocations grows from roughly 300 per month at the beginning of our observation period to a peak of 800–1000 per month at the start of 2011, after which it drops to around 500 per month in the last year, as the number of available addresses at RIRs has dwindled.[1] Overall, we find nearly 69K IPv4 prefix allocations at the beginning of our dataset and just over 136K at the end. This represents an increase of 67K prefixes—or, less than a doubling of the number of IPv4 prefix allocations over the course of the previous ten years. The figure contains a ratio line to show the relative allocation of IPv6 versus IPv4. We find that at the end of December 2013, on a monthly basis, the ratio of IPv6 to IPv4 prefix allocations is 0.57 and following a general upward trend. Thus, *although there are still significantly more allocated IPv4 prefixes (136k) than IPv6 prefixes (18k), and the monthly rate of IPv4 allocations is still about double, we see the IPv6 allocation rate continuing to grow while the IPv4 rate declines.* The ≈300 IPv6 versus ≈500 IPv4 allocations per month suggest *IPv6 is, for the first time, being deployed or planned on a majority of new networks.*

A2: Network Advertisement

Address allocation is a start, but to be used for Internet traffic IP addresses must be advertised in the global routing table. Therefore, our second metric is the number of IPv6 prefixes found in the Internet's global routing table. The Route Views project [38] and RIPE [32] both have a number of routers used for data collection, each peering with production Internet routers to obtain the routing tables from those peers. Based on routing table snapshots made available by these collection efforts, we obtain the number of prefixes announced on the first day of each month from January 2004 to January 2014. While these routing datasets are known to have biases, as we elaborate in § 6, these biases are not expected to affect the view of *globally-reachable* network prefixes.

Figure 2 shows the number of announced prefixes over time. We find 526 IPv6 prefixes on January 1, 2004. In January 2014, 19,278 IPv6 prefixes were advertised—an increase of 37-fold over

---

[1] We elide the April 2011 point such that the remainder of the plot is more readable. During that month, we find 2,217 IPv4 prefix allocations. This corresponds with APNIC's IPv4 pool dropping to a single remaining /8 and their "Final /8 Policy" being invoked, which caused a brief spike in allocated prefixes [3].

the course of ten years. For comparison, we also show the average number of IPv4 prefixes advertised per day; these increased fourfold from 153K in 2004 to 578K by 2014.

While total and monthly allocations and advertisements are both still higher for IPv4, the rate of IPv6 allocations is increasing at a faster pace than IPv4. This is expected since IPv4 has been an Internet reality for 30+ years now, and, hence, the need for additional addresses is, naturally, incremental. *The rate of IPv6 prefix allocations is now where IPv4 was 8 years ago.* What is more, in 2013 the monthly volume of allocations of IPv4 has dropped significantly, to 2009 levels, likely due to the exhaustion events starting in 2011. In sum, *the allocation and advertisement numbers and rates provide the basis for wide-scale Internet adoption of IPv6 from the addressing perspective.*

5. NAMING

Once IPv6 addresses are allocated and announced by routers, they must be used. The typical way addresses are referenced by Internet users and applications is via Domain Name System (DNS) names. Our next three metrics, therefore, focus on the prevalence of IPv6 support and use within the DNS ecosystem. A detailed description of DNS [28] is beyond the scope of this paper, but we remind the reader of some basic terminology. The authoritative groupings of names in the DNS hierarchy are called *zones*. DNS domain names map to IPv4 address via *A* records and to IPv6 addresses via *AAAA* ("quad a") records. DNS servers that manage zones and return records are called *authoritative nameservers*, while servers that execute queries on behalf of (usually many) users are broadly called *resolvers*.

N1: DNS Authoritative Nameservers

Our first naming metric aims to understand the prevalence of authoritative nameservers that themselves can communicate via IPv6. While IPv6-addressed nameservers are not required for an organization to employ dual-stack IPv6 (i.e., it could serve AAAA records via IPv4 nameservers), we believe that the prevalence of such nameservers offers telling evidence on the adoption of IPv6, especially by content providers.

The top level of DNS has been IPv6-enabled since 2008, when root nameservers deployed AAAA records [21]. As of Jan. 2014, reports from Hurricane Electric show that 91% of the 381 top-level domains (TLDs) also have IPv6-enabled nameservers [25]. These include the largest TLDs, such as, .com, .net, .cn, etc. Of the thirteen .com and .net nameservers, all can serve AAAA but only two (a. and bmgtld-servers.net) are themselves IPv6-addressable. To understand the prevalence of IPv6 nameservers for second-level domains (e.g., example.com), we survey the .com and .net TLD zones. We analyzed sample .com and .net zone files between April 2007 and January 2014 to track the prevalence of DNS glue records for authoritative nameservers in the zones.

Figure 3 shows the number of A and AAAA glue records in the .com and .net zones over the last 7 years. IPv6-enabled nameservers (AAAA records) are dwarfed by IPv4 nameservers (A records), but both show long-term growth. Following the pattern of other metrics, the growth rate (second derivative) of IPv6-capable nameservers is higher than that of IPv4, and the ratio of AAAA to A is increasing. As of January 1, 2014, the ratio of AAAA to A glue records for .com is 0.0029. We also show the AAAA to A ratio from Hurricane Electric's published probing data, starting in 2009, wherein A and AAAA lookups for all domains in the zone are periodically performed [25]. Few nameservers in general have glue records in their zone, and IPv6-enabled ones seem to have this configuration even less often. The ratio of domains actually returning AAAA records via queries (vs. A) is an order of magnitude higher (0.02 for .com) than the glue record ratio.

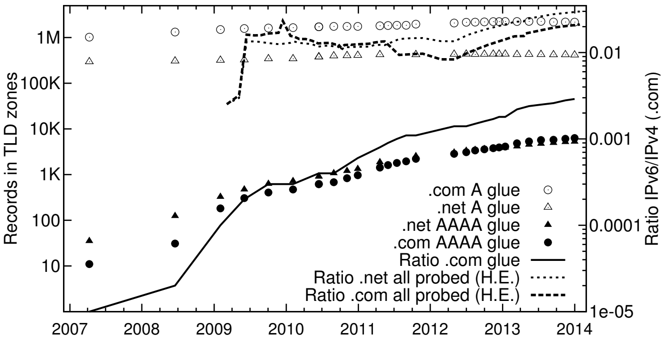

Figure 3: IPv6 nameserver and domain readiness. We see a steady increase in the glue records and a general increase in the probed domain names.

In sum, *the authoritative nameserver data indicates low (0.0028 for .com glue, 0.02 for probed) but increasing (56% growth in 2013 for glue) support for IPv6 in the overall .com and .net zones.*

N2: DNS Resolvers

A second naming metric we consider is the prevalence of resolvers requesting AAAA records. Due to caching within the DNS system, this is not a direct measure of demand; however, the number of resolvers looking up AAAA records indicates the breadth of the use of IPv6, and speaks to the basic capability of resolvers to issue AAAA queries as well as the existence of at least some clients within a resolver's pool making AAAA requests. Viewed over time, this can be used to gauge whether demand for IPv6 content is widespread or only from pockets of the network.

Packet Datasets for .com and .net: As an initial assessment, we examine two large datasets of packet-level DNS query traffic to the .com and .net TLD authoritative nameservers on five sample days between June 23, 2011 and December 23, 2013. One dataset consists of IPv4 packets, while the second contains IPv6 packets. Both are from Verisign, the registry operating the .com and .net zones. The IPv4 queries were captured at between three and five of the 17 largest globally-distributed .com and .net TLD server clusters (e.g., in Feb. 2013, from Dulles, VA; New York, NY; San Francisco, CA; and Amsterdam, NL). Our IPv4 data includes transactions with several instances of the lettered X.gtld-servers.net TLD nameservers. These 24-hour IPv4 datasets range from 2.3Bn to 4.2Bn queries (except for the first IPv4 sample, which only included 30 minutes of data and ≈110M queries). These same 17 global Verisign clusters support IPv6 traffic. The IPv6 samples analyzed were also each 24 hours and consisted of 420M–1,052M queries. While the packet collection apparatus for both datasets is known to be lossy, we performed analysis that suggests no systemic network effects that would skew the measurements we report [8].

We note that the IPv4 and IPv6 datasets shed light on slightly different aspects of adoption. The IPv4 data give us insight into behavior of networks that are not using IPv6 for their naming infrastructure but that happen to have clients and resolvers that make AAAA queries, which—given support by the resolver—is largely determined by operating system and application behavior. On the other hand, the IPv6 packet data represent networks where DNS resolvers are able to communicate via IPv6 to the .com and .net nameservers, which suggests a more advanced level of IPv6 adoption. Thus, the latter may be more representative of the behavior of fully-capable clients, whereas the former represents clients that

Table 3: Percentage of resolvers making AAAA queries to .com and .net. While under a third of all IPv4 resolvers (N=3.5M in latest sample) make AAAA queries, most of the IPv6 packet population of resolvers (N=68K) does.

Resolvers	2011-06-08	2012-02-23	2012-08-28	2013-02-26	2013-12-23
IPv4 All	33%	28%	26%	30%	31%
IPv4 Active	90%	93%	83%	93%	94%
IPv6 All	74%	77%	74%	82%	76%
IPv6 Active	99%	99%	99%	99%	99%

have software requesting AAAA records without the client necessarily having the ability to use them. For instance, Microsoft Windows XP clients that had a Teredo-configured [19] IPv6 addresses would, along with A, request AAAA records for names queried. However, these Teredo-based connections, either due to failure or preference, were found to be rarely completed by dual-stack hosts [41]. Windows Vista and later do not make AAAA queries when only a Teredo tunnel is available [9].

The resolver counts are, within an order of magnitude, stable over this period, with 3.5M seen in the most recent IPv4 sample and 68K in IPv6. Resolvers can service multiple, sometimes millions, of clients; so, this data represents the queries of many more than 3.5 million actual users. Although a single user or device can be configured to act as its own recursive resolver (e.g. by installing *bind*), we are more interested in the capabilities of resolvers serving multiple users. Therefore, in addition to aggregate results, we also report on a subset of the most active resolvers—e.g., enterprise or ISP-level—that send 10,000+ queries in a day.[2] There are 40K such active resolvers in the most recent IPv4 sample and 6K in IPv6.

In table 3 we show the percentage of resolvers in the two datasets that query for AAAA records. We see that nearly a third of all resolvers via IPv4 and three quarters via IPv6 make AAAA queries, as does the vast majority of active resolvers. Again, we stress this is not a measure of *use*, but an indication of *support* for IPv6 name resolution from within larger enterprises and networks. These numbers suggest that, *while AAAA records aren't in demand in every small corner of the network, at the organization or ISP level, IPv6 name resolution appears widely supported.*

N3: DNS Queries

In addition to the numbers of IPv6-addressable nameservers and resolvers requesting IPv6 addresses measured above, a final naming component we consider is the distribution of actual IPv6-related DNS queries. This speaks to *how* naming is being used in IPv6. We first determine whether IPv6 users are interested in the same names as IPv4 users. This will inherently be influenced by user population differences (e.g., regional and sample effects), including client OS differences (which construct DNS requests). Therefore, differences are expected. To measure the agreement between queried domains via A and AAAA records in our two .com/.net packet samples, we calculated Spearman's rank correlation coefficient (ρ) between the top 100K domains by each of the four types (IPv4 sample A and AAAA, and IPv6 sample A and AAAA). We limited analysis to the most-queried 100K domains in order to avoid skewing results by rarely-queried domains, such as typos, but we wanted a large number in order to capture a diverse set of content.

Table 4 shows the results. As a preface, rank correlation is, by definition, lower than set intersection, and the intersection numbers (not shown) for the three sets of domain list pairs range from 55% to 84%. We see that the domain rank correlations between the IPv6 and IPv4 samples via the same record types are moderate to strong ($\rho \approx 0.70$), indicating that *domain interest is similar between users of IPv4 naming infrastructure and those using IPv6.* Still, differences remain, and no clear trend is visible in this time period. Likewise, when we examine the within-packet-sample cross-type correlations (e.g., A vs. AAAA for IPv4), we see much less correlation. We suspect this is in part due to the fact that, whether an A or AAAA is requested by a given host is determined by applications and OSes in use, but there is greater similarity when examining the same application/OS patterns across protocol packet samples. Some of these differences may be accounted for by the differences in set sizes. Across the five samples, the median percentage of queries that the top 100K domains account for is 55% for A via IPv4 and 60% for A via IPv6; for AAAA, it is 77% for IPv4 and 42% for IPv6. No clear trend is evident. In sum, the data suggests *differences in application use of IPv6 (which is expected due to the N1 results) but marked overlap in the domains of interest to networks using IPv6 resolvers versus those using IPv4..*

Table 4: Spearman's ρ rank correlations for top 100K domains queried by A and by AAAA via IPv4 and IPv6 ($P < 0.0001$ in all cases). There is moderate to strong (darker grays) correlation between IPv4 and IPv6 domains for same record types.

Domain Lists	2011-06-08	2012-02-23	2012-08-28	2013-02-26	2013-12-23
4.A : 6.A	0.65	0.73	0.70	0.70	0.57
4.AAAA : 6.AAAA	0.69	0.80	0.82	0.74	0.68
4.A : 4.AAAA	0.32	0.32	0.35	0.34	0.42
6.A : 6.AAAA	0.29	0.23	0.20	0.26	0.32

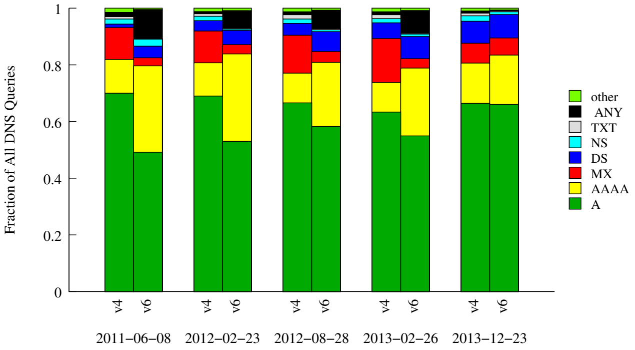

Figure 4: Breakdown of query types across five IPv4 and IPv6 DNS samples between Jun. 2011 and Dec. 2013 (key is in stack order). The distribution of IPv4 and IPv6 query types within each day converges over time (p<0.05).

Turning from the names in the queries to the records, Figure 4 shows the top seven record types (plus all others under the "other" category) requested in the IPv4 and IPv6 packets on the five days of samples. We observe that, while there are still some differences in the distributions, *there is a statistically-significant convergence of query types over time (average monthly difference decrease of 1.65% with p<0.05), and the query types in IPv6 are now much more similar to IPv4 than just two and a half years ago.*

6. ROUTING

Once IPv6 addresses are allocated and advertised, as well as potentially being named, the next prerequisite for using the new protocol is routing. While routing itself has many components, and we have already discussed IPv6 prefix advertisement in section A2, a key aspect of routing that deserves careful measurement is topology. The richness of the IPv6 topology, in terms of the number

---

[2] This threshold is arbitrary. We certainly miss smaller organization-level resolvers, but the included ones are very active.

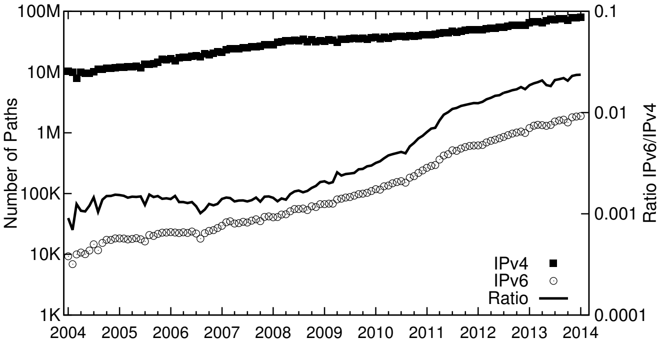

Figure 5: Number of globally-seen IPv4 and IPv6 paths. There is a 110-fold increase in IPv6 paths over ten years.

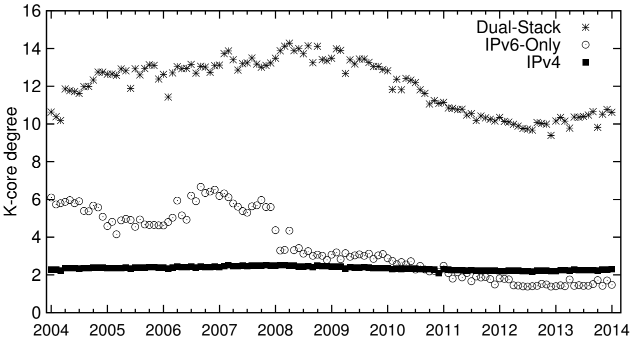

Figure 6: AS centrality. Pure-IPv6 and dual-stack ASes are becoming more prevalent at the edge.

and length of paths and the connectivity of ASes, speaks to the resilience, and, thus, production-readiness of the network.

T1: Topology

The IPv4 topology has been studied in depth (e.g., [26, 42]), but we also need to understand the relationships between organizations with respect to external IPv6 routing capability and connectivity to understand the overall strength or brittleness of the network. As we did for the advertisement metric (A2) we use all of the routing table snapshots collected by Route Views [38] and RIPE-NCC RIS [33] between Jan. 2004 and Jan. 2014 in the following.

Routing Table View Biases: Before we delve into the insights that these data afford, it is important to understand possible bias. As noted by earlier studies (e.g., [15]) the global public routing datasets available (e.g., Route Views, RIS) suffer from at least two forms of bias. The first is geographic, in that global routing data is collected from a finite set of collectors, whose global distribution is not uniform, leading, for example, to fewer samples from the African continent. The second bias stems from most of the data in these collections coming from volunteer networks that turn out to generally be large top-tier ISPs. Therefore, many peer-to-peer paths between smaller ISPs are not visible in the data, since these routes are never propagated to the top-tier ISPs. These biases are a limitation of the data. However, we believe that even though imperfect the data still yields useful information about IPv6 adoption for the following reasons: (*i*) no substitute data lacking bias exists (e.g., traceroute also has bias); (*ii*) the view of the global routing infrastructure of ISPs whose routing data is represented is a real view from their perspective, and any path or AS counts observed are, at worst, lower bounds; (*iii*) we have no reason to believe that the bias present in these data for IPv6 differs systematically *relative* to that for IPv4, suggesting that looking at ratios of adoption, especially over time, is reasonable; and (*iv*) in the cases of counting globally-visible prefixes or ASes seen supporting IPv6, the fact that some local paths are missed does not speak to the *global* adoption state. Therefore, we present the data knowing full well it is a less-than-full statement on the routing state. We encourage the community to collect and refine our analysis using better data.

AS Support and Connectivity: We first examine the number of ASes supporting IPv6 globally as well as the number of unique globally-visible AS-paths (i.e., the paths with unique AS sequence). Both are indicators of IPv6 adoption, mostly at the service provider level. AS adoption is indicative of *support* for IPv6, while the number of AS-paths is an indicator of maturing *connectivity* between ASes. We omit the figure showing AS-level adoption in favor of Figure 5, which shows the number of unique IPv6 and IPv4 paths announced on the first day of each month. We observe that the number of IPv6 paths has a 110-fold increase from January 2004 to January 2014, while there is only an eight-fold increase in the number of IPv4 paths. However, the IPv6 to IPv4 ratio is only 0.02 in January 2014, indicating the IPv6 routing mesh is still at an early stage of maturity. AS-level support for IPv6 is not shown, but follows a faster upward trend, with an 18-fold increase (versus two-fold for IPv4) and the current ratio of IPv6 to IPv4 ASes is 0.19 – almost ten times the path count ratio. As expected, the indicator of ASes *supporting* IPv6 leads the measure of *connectivity*. Again, we note that, because of possible bias in the view afforded by this data, raw numbers of the IPv4 and IPv6 paths seen are less meaningful than is the overall ratio.

AS Centrality: To understand the topological position of IPv6 ASes, we next compute the *k-core degree* of each AS in the topology graph. A k-core of a graph is the maximal subgraph in which every node has at least degree k. A node has k-core degree of N if it belongs to the N-core but not to the $(N+1)$-core. As used in [17], this measure represents a natural notion of the *centrality* of ASes. In other words, ASes with a high k-core represent well-connected, typically large, ISPs, while those with low k-core represent edge or stub networks. We show the average k-core degree of ASes in Figure 6. We find that dual-stack ASes have a much higher degree of centrality than other ASes. In 2004, the pure IPv6 ASes were located in a relatively central position. However, we see pure-IPv6 ASes, a small fraction of all, becoming more prevalent at the edge after 2008. Our results are in accordance with those of CAIDA [10], who report that IPv6 is largely deployed at the core but lags in edge networks. Note [10] uses a deeper and more robust analysis of these same public datasets, wherein, notably, they filter out transient links. The numbers we find indicate *dual-stack becoming more widely deployed among well-connected central ISPs, while single-protocol networks are mostly those at the edge. In other words, the older edge networks are the laggards.*

We caution that studying native IPv6 topology is useful but insufficient. Transition from IPv4 to IPv6 introduces a co-dependence between the protocols. Therefore, unlike when studying IPv4 topology independently, when studying IPv6, we must consider the parts of IPv4 that glue together "islands" of IPv6. An in-depth analysis is beyond scope, but we point readers to recent work in [10].

7. END-TO-END REACHABILITY

Having dealt with the prerequisites of addressing, routing, and naming in the previous three sections, we now turn to the readiness of Internet end hosts to use IPv6. We split this into two metrics for the readiness of service-level devices and client-level devices.

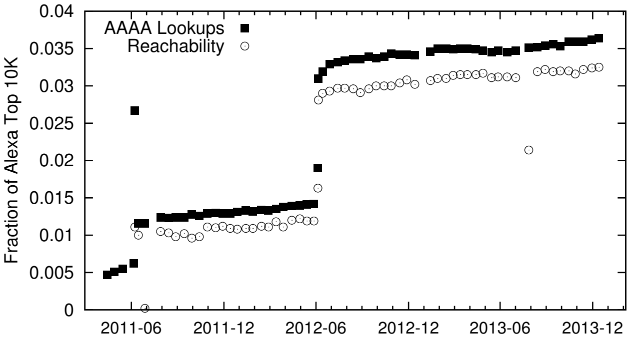

Figure 7: Fraction of top 10K sites with AAAA records and reachable via IPv6. Two discontinuous jumps correspond to World IPv6 Day 2011 and World IPv6 Launch 2012.

R1: Server-Side Readiness

Obviously, wide-scale adoption requires services to be capable of handling IPv6 traffic; therefore, our first approach to end-host readiness involves assessing prevalence of IPv6-enabled services.

While not indicative of all services, one way to assess IPv6 service penetration is to measure popular web servers. Much like Nikkhah et al [29] and with congruent results we use Alexa [1] to determine the most popular web sites. We then determine which sites have a AAAA record in DNS, and, for those that do, we then test reachability of the web site via a tunnel to Hurricane Electric. Ideally, this metric tries to assess the server, but we have no way to do so without also assessing the path to the server. Hence, our measurements offer an approximation. We have been probing the top 10K web sites for AAAA records since April 2011 and for reachability since June 2011.[3] Figure 7 shows our results. We first note a jump in June 2011 that corresponds to World IPv6 Day. We find a roughly five-fold increase in AAAA records available at that point. However, we also see a nearly immediate fallback. This is understandable given that the stated goal of that day was merely to serve as a "test flight" of IPv6 capabilities, rather than to permanently enable IPv6 services [22]. Subsequent to this drop off, in spite of the limited goal, we find that World IPv6 Day 2011 is responsible for a sustained two-fold increase in the IPv6-capable web sites. In the following year, the June 2012 World IPv6 Launch Day also resulted in a sustained doubling of AAAA records. Further, aside from the two jumps, *we find a slowly growing trend across time with over 3.2% of the Alexa top 10K now being reachable via IPv6.* It is notable what an impact concerted community efforts, such as the two IPv6 readiness/launch days can have on IPv6 server readiness.

The second set of points on the plot show reachability. The data shows that most of the hosts for which we find AAAA records are also reachable. Further, the reachability trends generally mirror those for web servers having AAAA records. Our results generally agree with [29]. In conclusion, *while only about 3.5% of the top most popular websites are IPv6-ready, there has been significant growth in the last three years, and large jumps are possible.*

R2: Client-Side Readiness

In addition to IPv6-capable services, clients need to be IPv6-enabled as well. This metric examines the ability of client systems to employ IPv6, subsuming all that is required on the client side (i.e., working IPv6 network transport, DNS, operating system, etc.).

Google makes aggregate data about client adoption of IPv6 available on an ongoing basis [14]. Their experiment consists of adding a JavaScript applet to search results from `www.google.com` for a randomly sampled set of users [29]. The script first performs a name lookup on one of two experimental host names and then sends a request to the IP address returned in the DNS response. In 90% of the cases the script chooses a name representing a dual-stacked server, while in the remaining cases a name representing a IPv4-only server is chosen for comparison purposes. The addresses point to 2–5 data centers (in Asia, the US, and Europe). The experiment is conducted millions of times per day. Note that, as with the R1 measurements, this data again conflates the client capabilities with those of the path from the client, and, therefore, this is an approximation of the ideal metric.

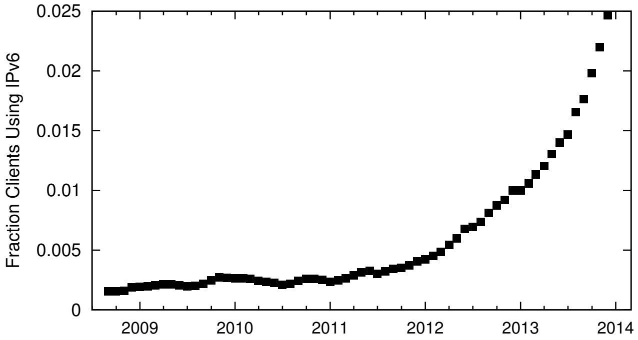

Figure 8: Average monthly fraction of clients able to access Google over IPv6. 2.5% of clients use IPv6, but this number has been growing sharply. The most recent two-year annual growth rate averages 150%, a more than doubling each year.

Figure 8 shows the average monthly fraction of clients that connect to Google via IPv6 over the last 5+ years. The plot shows a growth factor of 16 over the course of the dataset—from 0.15% in September 2008 to 2.5% in December 2013. Further, most of the growth comes in the last two years, where the ratio increased markedly, by 125% in 2012 and 175% in 2013, more than doubling each year. As discussed in section A3, this measure is probably somewhat optimistic; since Google has many direct private peerings to ISPs, some clients may be able to reach Google by IPv6 but not other content. However, these numbers are roughly in line with those reported in another large client study [41], which found that although 6% of a global sample of clients were IPv6-capable, only 1-2% of dual-stack preferred IPv6. In sum, *the data shows very strong growth in Google clients' ability to use IPv6, especially in the last two years.*

8. IPV6 USAGE PROFILE

While the metrics and data sketched in the previous sections set the stage for IPv6 adoption by measuring addressing, routing, naming, and end-host capabilities, in this section we aim to directly assess IPv6 traffic "in the wild". That is, we aim to understand the operational characteristics, or usage profile, of how IPv6 is actually employed by those that have adopted it.

U1: Traffic Volume

Our first traffic-related metric simply aims to understand how much of Internet traffic is using IPv6. We begin by introducing the Internet traffic datasets we contribute.

Arbor Internet Traffic Data: We assembled two datasets describing the traffic traversing customer networks monitored by devices from Arbor Networks, a provider of traffic analytics and security devices for large networks. Arbor's customers include a signifi-

---

[3] Our probing data is available [2].

cant number of large ISPs in the world, and, in aggregate, they have visibility into nearly half of Interdomain (i.e., Internet) traffic. These two datasets consist of traffic summaries (daily netflow statistic aggregates, including protocol, port, and volume information), from peering, aggregation, and customer-facing routers at participating networks. The first, older, dataset tracked daily peak five-minute traffic volume for both IPv6 and IPv4, measured a sample of 12 Arbor customers providing anonymous data, and included data from the second quarter of 2010 through February 2013. The second Arbor dataset (for 2013) reports daily average volume, includes approximately 260 providers of anonymous data that together represent an *estimated 33-50% of global Internet traffic*, and collects data using the same methodology as Labovitz et al. [24] (now with more providers). We include the larger dataset starting in January 2013; however, even the smaller, 12-provider sample covers an aggregate of over 400 routers and 55K links, representing a cross-section of different-mission and varying-size Internet organizations. Both samples include global Tier 1 ISPs, national and regional Tier 2 ISPs, content/hosting providers, and universities. In the newer dataset, which represents 19 Tier-1 and 92 Tier-2 providers plus over 100 enterprises, content providers, etc., each continent is represented, as are both fixed-line and mobile Internet providers. We refer to the older (smaller) and newer (larger) datasets as $\mathcal{A}$, and $\mathcal{B}$, respectively. In the fourth quarter of 2013, the daily median Internet traffic in dataset $\mathcal{B}$ was 58 Tbps.

We normalized the traffic measurements by the number of Arbor Networks providers in the samples, to distinguish organic traffic growth from growth due to changes in the number of customers. Figure 9 shows the median daily peak and average traffic volume for each month in our two datasets, respectively. Since one set of points and line represent 5-minute peaks, and the other averages, we present the data as separate points. This difference helps explain the shift between the lines visible for the two months for which we had both datasets in January and February, 2013.

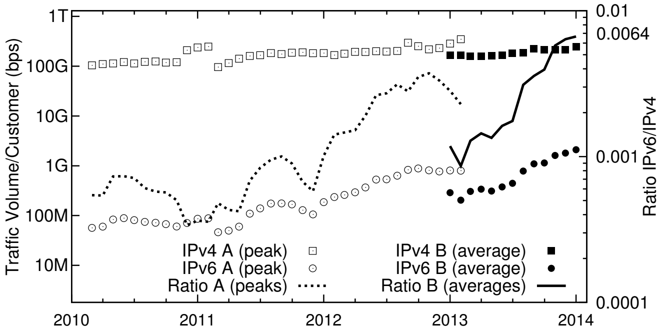

Figure 9: Global Internet traffic data in two datasets: $\mathcal{A}$ for Mar. 2010–Feb. 2013, monthly median *peak* 5-minute volume for 12-provider sample; and $\mathcal{B}$ for 2013, the monthly median of daily *average* traffic volume for ≈260 providers. IPv6 is 0.6% of traffic, and 2-year growth relative to IPv4 is 451% annually.

The figure shows that IPv6 is still dwarfed by IPv4 traffic (by roughly two orders of magnitude). However, both IPv4 and IPv6 peak traffic volumes are generally increasing, and IPv6 is on a strong upward trajectory. Over our measurement period we find roughly an order of magnitude increase in the median daily peak volume for both protocols. As the ratio line (representing the raw traffic, not normalized by customers) shows, we do find a significant rise in IPv6's relative contribution to Internet traffic. In March of 2010, the ratio of IPv6 to IPv4 is 0.0005, while in December 2013 it is 0.0064—a 13-fold increase. In sum, *while, overall, the proportion of IPv6 traffic on the Internet is still under one percent, it has grown, relative to IPv4, by 433% percent year-over-year in 2013, 469% in 2012, and 71% in 2011, a rapid pace.*

U2: Application Mix

Another important metric when studying the IPv6 usage profile is what applications are used. This can, for instance, inform our understanding as to whether IPv6 is being used for typical user activity or for specialized use, as was reported in the past (e.g., [23]).

Table 5: Application mix (%) between Dec. 2010 and 2013. HTTP/S increases from 6% to 95% for IPv6, surpassing IPv4, while back-end services (e.g., dns, ssh, rsync) decline significantly. IPv6 usage looks much more like IPv4 than in the past.

Application	Dec 2010	Apr/May 2011	Apr/May 2012		Apr–Dec 2013	
	IPv6	IPv6	IPv6	IPv4	IPv6	IPv4
HTTP	5.61	11.81	63.04	62.40	82.56	60.61
HTTPS	0.15	0.88	0.39	3.91	12.66	8.59
DNS	4.75	9.11	4.09	0.14	0.33	0.22
SSH	0.56	3.73	2.65	0.11	0.27	0.20
Rsync	20.78	5.11	2.65	0.00	0.13	0.00
NNTP	27.65	5.84	1.03	0.13	0.00	0.25
RTMP	0.00	0.05	0.11	2.39	0.00	2.74
Other TCP	*	*	18.72	3.20	1.66	4.08
Other UDP	*	*	1.73	11.90	0.27	2.82
Non-TCP/UDP	*	*	4.94	14.10	2.11	20.21

We have application information from Arbor Networks for the same traffic samples described earlier (although we are missing IPv4 data prior to 2012). The network flow monitors classify traffic by port number, and, hence, the categorization may not be completely accurate. For instance, we note that HTTP port 80 is often used for tunneling non-web applications, as it tends to be open in firewalls. However, we believe this categorization is useful as a first order analysis. Table 5 shows the proportion of traffic for each application that makes up at least 1% of either IPv6 or IPv4 traffic. We see in the 2012 and 2013 samples that HTTP dominates within both IPv6 and IPv4. In the 2013 sample, HTTPS has increased significantly in IPv6, surpassing even that in IPv4. Likewise, we see a dramatic uptick in HTTP/S traffic across the four years. The large fraction of likely web traffic now observed, at 95%, is a major departure from the patterns observed in earlier studies of IPv6 dating from 2008 and early 2009, which report HTTP traffic volume below one percent [23, 36]. Interestingly, Karpilovsky [23], Savola [35] and Hei [18] also report large amounts of DNS traffic (e.g., 80-90% in 2008 [23]), which continued to rank highly in our own data until 2012. Our 2013 sample is the first time we see DNS decreasing to IPv4 levels and actual content (HTTP/S) surpassing even what is seen for IPv4; this is a significant evolution. We note that one possible factor at play in the reduction of DNS traffic is the behavior of newer Windows operating systems (starting with Vista), which, as mentioned earlier, no longer make AAAA queries when their only IPv6 connection is via Teredo [9]. Likewise, we see a substantial reduction in IPv6 NNTP traffic [11]. We believe the previously large volume seen was related to (*i*) the fact that a large NNTP path is part of the traffic sample and (*ii*) NNTP was heavily used for piracy over IPv6 when several USENET services that otherwise required paid subscriptions offered free IPv6 access[4].

Additionally, in the 2013 sample, we find 4.04% and 27.11% of the traffic volume not ascribed to a particular application for IPv6 and IPv4, respectively, a large decline from 2012 for IPv6 and a smaller one for IPv4. However, the distribution of non-identified traffic is different between IPv6 and IPv4. For example, while most

---

[4] e.g., http://www.techjawa.com/2011/02/10/guide-get-free-usenet-access-with-ipv6/

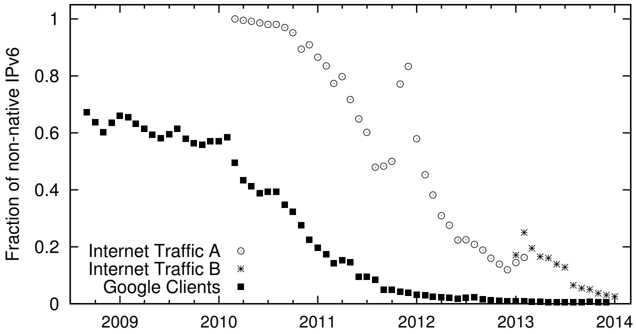

Figure 10: Fraction of IPv6 traffic carried by the two most prevalent transition technologies in the Internet traffic and Google client samples. Non-native IPv6 traffic was the majority of packets 3-4 years ago, but currently represents less than three percent of traffic and one percent of Google clients.

of the bytes in IPv4 are non-TCP/UDP at 20.21%, such traffic only contributes 2.11% of the overall bytes in IPv6. Although we were unable to investigate this "other" category more deeply, we speculate that the usage of peer-to-peer and similar popular non-well-known-port applications still differs between IPv4 and IPv6 in the TCP/UDP categories, while ICMP and tunneling protocol mix differ in the non-TCP/UDP category. The method of aggregation of untracked protocols for 2010 and 2011 does not allow comparison, unfortunately. To summarize, *over the last three years we see a dramatic evolution in IPv6 application use, wherein content packets now far outnumber infrastructure service packets (DNS, ICMP), and the profile now resembles IPv4.*

U3: Transition Technologies

IPv4 and IPv6 coexistence is greatly complicated by the lack of backward compatibility. In what is now acknowledged as one of the most significant IPv6 design limitations, native IPv6 network devices cannot communicate with their IPv4 counterparts without an explicit network translation layer [27]. As a result, the success of any large-scale IPv6 transition depends on the complex interplay between the cost and scalability of translation technologies and the commercial incentives (or disincentives) motivating the transition to native IPv6 infrastructure. A common transition technology is tunneling. Tunneling technologies interconnect "islands" of IPv6 using encapsulation across IPv4 infrastructure, or vice versa. In addition to tunnels, Teredo [19] provides IPv6 connectivity to hosts behind IPv4-NATs using UDP-encapsulation. Our next metric aims to understand the prevalence of the most common transition technologies being used in the wild where end-to-end IPv6 addressing is not fully in place. As IPv6 matures, we expect a smaller fraction of its traffic to use these technologies and more of it to be native.

Both the Google client and Arbor Networks datasets described earlier include information on the prevalence of various transition technologies. The Google perspective provides a view on the capabilities of end hosts, while the Arbor view is an assessment of actual Internet traffic. Figure 10 shows the prevalence of non-native IPv6—which is defined as Teredo and IP protocol 41 traffic (used by 6to4 and 6in4). The two Internet traffic data points $\mathcal{A}$ and $\mathcal{B}$, refer to the Arbor Internet traffic datasets, $\mathcal{A}$ and $\mathcal{B}$, described earlier. The Google data shows that, while in 2008 only 30% of IPv6-enabled client end-hosts could use native IPv6, that number has increased to above 99% over the last four and a half years.

In 2010 we find the Internet traffic data shows nearly all IPv6 traffic using some tunneling technology. However, as of the end of December 2013, nearly 97% of the traffic is now native. We note that, of the tunneled IPv6 traffic in late 2013, IP protocol 41 dominates, contributing over 90% of the tunneled volume compared to less than 10% for Teredo. The Arbor numbers between mid-2011 and February 2012 correspond roughly to earlier measurements from that time, (e.g. [34] and [41]), while the Google numbers show much less transition technology used than those studies; this may be explained by the direct peerings phenomenon, described above. Overall, the data shows that *native traffic is now the vast majority of IPv6 traffic, a dramatic change from just three years ago. The Internet's IPv6 traffic is now real IPv6.*

9. PERFORMANCE

A crucial metric of IPv6 adoption is performance—which can mean different things depending on the measurement perspective we take (e.g., the speed of a site to load for a user, or, for an ISP, the bandwidth across peering links). Several works predating IPv4 address exhaustion offer initial results in the area of IPv6 performance (e.g., [5, 7, 39, 44]). Further, there are some performance results since the exhaustion milestones [10, 29]; both of these latter studies report that performance over IPv6 paths that align with IPv4 at the AS-level is similar for the two protocols but differs when paths diverge. More recent work exploring a methodology for passive measurement of IPv6 and IPv4 performance was contributed by Plonka and Barford [31], who found great variability in relative performance of the protocols in a campus traffic sample. The data we present here aims for a global and longitudinal, if less granular, examination of relative IPv6 network performance.

P1: Network RTT

We suspect that hardware, software, and configuration differences could result in different quality of data transmission in IPv6 versus IPv4; indeed, previous measurements and ones we report here have shown differences. Although actual client-to-service network performance for large global sets of clients and services would be a more ideal metric, we use the approximation of average 10- and 20-hop round trip time (RTT), tested from dozens of global perspectives, as a proxy. This hides the details of path differences and heterogeneity of end points, allowing a simple apples-to-apples comparison of raw network performance over the same number of IPv4 versus IPv6 nodes. We don't claim that this is the only or the best IPv6 network performance metric, but it serves as a reasonable approximation of long-term performance evolution.

Our analysis is conducted on the traceroute-based performance data collected by CAIDA Archipelago Measurement Infrastructure (Ark) [4] to measure RTT in IPv4 and IPv6. Globally-distributed Ark monitors probe all IPv4 /24s and all announced IPv6 prefixes continuously. We analyze data from December 2008 to December 2013. While this dataset is also the basis of earlier work ([10]), we re-analyze an updated longitudinal version to observe performance in the context of the other metrics we report.

Figure 11 shows the median RTT with hop distances 10 and 20 for each month. We find that in 2009 RTTs were roughly 1.5 times longer for IPv6 than for IPv4. While the IPv4 RTTs have increased slightly over this time period, IPv6 RTTs have decreased slightly. In 2013, the RTT for hop distance of 10 is almost identical for IPv4 and IPv6. IPv6 had better RTTs than IPv4 at 20 hops in 2012 through mid-2013. To compare relative performance, we show the IPv6 to IPv4 ratio for the 10-hop RTT, as it has been less favorable for IPv6. Since, the better the performance the smaller the RTT, we show the ratio of the reciprocal of RTT for each protocol. As noted in [10], the sample of IPv6 data is small and the results might be dominated by a few paths. Also, evolving tunnel use likely impacts

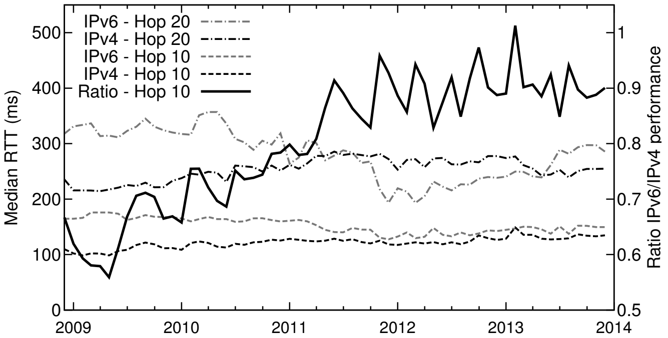

Figure 11: Median Round Trip Time (ms) with hop distance 10 and 20 for IPv4 and IPv6. IPv6 showed poorer performance before 2010, but the last several years have seen performance converging to within about .90–.95 of IPv4.

RTT and hop count. Thus, we cannot conclude that IPv6 has better RTT performance than IPv4, overall. However, *the long-term trend shows clear improvement for IPv6, and it has approached parity with IPv4 (≈95%), for the first time, in the last several years.*

10. IPV6 PRESENT AND FUTURE

We examined a set of twelve metrics for assessing the adoption of IPv6 based on a comprehensive set of ten longitudinal datasets—some original, some publicly available, most large and global. In some cases, we updated and replicated similar measures reported in years past (e.g. RIR allocation data, performance data); in others, we presented new large data samples (e.g. the Internet traffic data, the Verisign IPv6 .com and .net DNS packet data). Here, we first highlight what the current state of adoption looks like, when examined through the broad set of perspectives we consulted. After that, we provide rough estimates of where we expect adoption to be in five years, based on recent trends.

10.1 IPv6 Present

The Value of a Broad Approach: In Figure 13 we show five-year ratios of IPv6 relative to IPv4 for seven of our metrics. Most prominent is the result that different metrics give entirely different insight into the adoption of IPv6, and suggest orders of magnitude different progress. For instance, while roughly 36% of new monthly (and 12% of cumulative) allocated prefixes are IPv6, we find just 0.63% of average traffic is carried over IPv6—a two-order-of-magnitude difference. These differences across metrics serve to highlight that multiple viewpoints must be considered to fully understand the progression and true state of adoption. In addition to the differences seen when examining different types of data—i.e., different prerequisites or operational characteristics—differences within the same type, but from distinct perspectives, are also important to consider. For example, recall that the difference in non-native IPv6 traffic as seen by Arbor Networks versus that seen by Google in metric U3 has been noticeable. What is more, *the order of adoption, as reflected by the relative rank of metrics, generally follows the prerequisites for IPv6 deployment* (e.g., allocation precedes routing, which precedes clients, which precedes actual traffic).

IPv6 is Now Real: Compared to prior work and earlier data, *we see in our recent data a dramatic qualitative and quantitative evolution in the state of IPv6 adoption, indicating a major shift in how the protocol is being used in the last three years.* Table 6 summarizes the usage profile of IPv6, and how it has evolved over this time. Traffic data shows that IPv6, while just 0.63% of measured Internet packets, is growing at a rate of over 400% in each of the last two years; application mix data shows content packets now dominating traffic; transition technology data shows that virtually all IPv6 traffic and Google clients are now native; and, performance data shows IPv6 now nearly on par with IPv4. *IPv6 is finally being used natively, for production, and at a rapidly-increasing rate.*

Table 6: Measures of actual operational characteristics of IPv6, recently and three years ago. These suggest that IPv6 is now mature. We contend that IPv6, as a real, production protocol, has finally come of age.

Metric: Operational Aspect Measured	IPv6 Status at End of:	
	2010	2013
U1: IPv6 Percent of Internet Traffic	0.03%	0.64%
U1: 1-yr. Growth vs. IPv4 (*Mar-2010 – Mar-2011)	−12%*	+433%
U2: Content's Portion of Traffic (HTTP+HTTPS)	6%	95%
U3: Native IPv6 Packets vs. All IPv6	9%	97%
U3: Native IPv6 Google Clients	78%	99%
P1: Performance: 10-hop RTT^{-1} vs. IPv4	75%	95%

Inter and Intra-Regional Differences: Our cross-metric analysis allows us to see stark regional differences in adoption. For example, when breaking down the cumulative allocation data by RIR (A1), we find RIPE responsible for 46% of allocations, while ARIN is responsible for 21%, and APNIC 18%. These three RIRs represent the most-connected portions of the Internet, and, therefore, it is not surprising that they allocate most of the new prefixes. We also observe that LACNIC and AFRINIC are responsible for 12% and 2% of the allocations, respectively. However, although the well-connected regions dominate the absolute number of prefixes, the *ratio* of IPv6 to IPv4 prefix allocation per region tells a slightly different story. Here, we find that LACNIC has, by far, the largest ratio at 0.280, followed by RIPE at 0.162, AFRINIC at 0.157, APNIC with 0.143, and we see only half as much IPv6 prefix allocation, 0.072, for ARIN. Part of the likely reason is that the ARIN region was an early adopter of IPv4, accumulating many prefixes before resources became constrained.

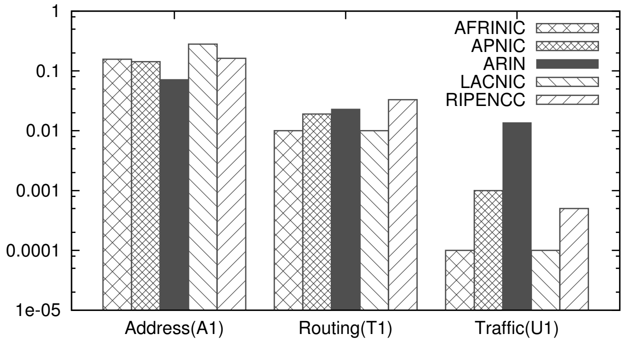

Figure 12: IPv6 to IPv4 ratio for three metrics, broken down by region. We see that, not only do different regions have different levels of IPv6 adoption, but that the level of adoption varies by layer; i.e., the relative rank of regions differs across metrics.

In Figure 12, we show an analysis of the IPv6 to IPv4 ratio for allocation (A1) as well as two additional metrics whose data allowed region differentiation (T1 [announced AS paths] and U1 [average traffic, $\mathcal{B}$]). Adoption level varies considerably across metrics for the regions (note the log scale), with the highest measured region for each metric at least three times higher than the lowest. We do not show its numbers, but the A2 metric ranks closely match T1, as they both track routing. While we might expect different

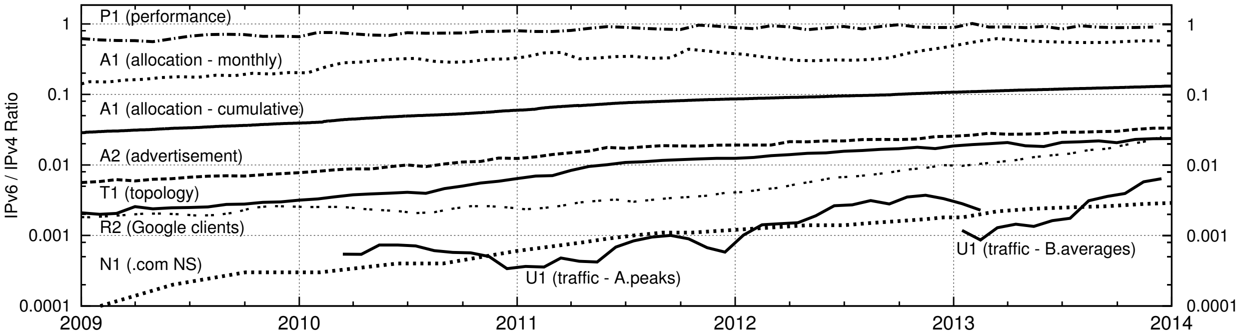

Figure 13: The ratio of IPv6 to IPv4 for seven metrics over the last five years, showing adoption level ranges by two orders of magnitude depending on metric, and accelerated growth. For instance, the two traffic lines, $\mathcal{A}$ (peak, through 2013-02) and $\mathcal{B}$ (average, for 2013), show IPv6 traffic has increased over 400% in each of the last two years.

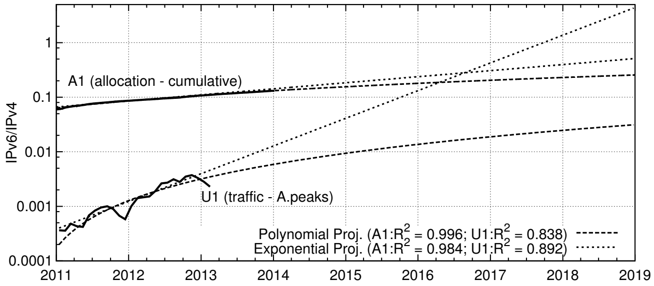

Figure 14: Trends for Allocation and Traffic (using the older $\mathcal{A}$ [peak] traffic data), starting with 2011, when IPv4 exhaustion pressure increased, and five-year projections. We caution that trends are volatile and prediction is hard.

regions to adopt IPv6 at different rates, surprisingly, we also see that the same ordering of regions does not persist across metrics. For example, the Address Allocation metric shows that ARIN lags behind. However, ARIN performs much better on the other two metrics. First, this affirms our argument that a single metric cannot fairly reflect IPv6 adoption status. More surprisingly, this suggests that, *not only are different regions adopting IPv6 at different rates, but there are different incentives and obstacles (perhaps resource constraints, policy, etc.) that vary the rate for different layers of adoption (i.e., metrics) in any given region.* We intend to explore the cause of these varying pressures in future work.

10.2 IPv6 Future

We close by attempting to model how future IPv6 adoption may proceed. As a baseline, for the metrics for which we have four or more years of data, we have already seen that, over that time, nearly every measure of adoption of IPv6 relative to IPv4 has increased by an order of magnitude. In Figure 14 we show the IPv6 to IPv4 ratio for A1: Address Allocation (cumulative) and U1: Traffic ($\mathcal{A}$, peak), between 2011 and 2013 or 2014. We start with 2011, as this is when IPv4 exhaustion pressure became more acute and we began to see larger increases in several of our metrics, including traffic. We chose A1 and U1 as they bookend the gamut of adoption metrics, showing the highest and lowest level, respectively. These also represent the first and last step in deployment. We use the older traffic sample, which ends in 2013, instead of the newer, as the former is for a longer period (but is more conservative than the newer, whose rate of increase in 2013 was 433%). The figure also includes projections out to 2019 based on both polynomial and exponential fit functions (R^2 values as shown). Of course, it is possible that upcoming IPv4 exhaustion milestones or other events will lead to discontinuities in the adoption trends. Additionally the growth rate may shift for a number of reasons. Thus, we caution that, while these models represent predictions for where IPv6 adoption may be in five years based on recent trends, the same trends have been volatile in the past and even a small shift could result in much different outcomes. With those caveats in mind, according to the model's projections, *we expect that by 2019 the number of IPv6 prefixes allocated will be about .25–.50 of IPv4, while the IPv6 to IPv4 traffic ratio will be somewhere between .03 and 5.0. In other words, IPv6 appears headed to be a significant fraction of traffic.*

11. LIMITATIONS AND FUTURE WORK

In addition to the stated limitations of our datasets, we admit that our framework and proposed metrics are missing some notable perspectives. Social, behavioral, and economic factors are worthy of study, and there are also other technical aspects we did not look at. For instance, vendor support, including in software (e.g., operating system) and hardware (e.g., routers) is useful to understand. Characterizing the prevalence and motivations of actors that forego adopting IPv6 in favor of alternatives, such as carrier-grade NAT (CGN), is also a valuable tangential perspective on IPv6 deployment. Further, examining the factors that lead to differing incentives and obstacles to adoption across and within global regions would be a fascinating and useful endeavor. Even without such broadening of perspectives, the overall topic of IPv6 adoption is too large for any of us to tackle alone; we invite the community to continue contributing measurements and hope our own data and meta-analyses add to the community's understanding of the Internet's largest transition.

Acknowledgments

This work was supported in part by the Department of Homeland Security Science and Technology Directorate under contract numbers FA8750-12-2-0314, and FA8750-12-2-0235; the National Science Foundation under contract numbers: (each CNS-) 0751116, 1111672, 1111699, 1213157, 1255153, 1330142, 08311174; and the Department of the Navy under contract N000.14-09-1-1042. The authors would like to also thank Arbor Networks, CAIDA, Google, RIPE-NCC, Route Views, and Verisign for making data available. We also thank Duane Wessels and Jared Mauch for supplying older zone files. Finally, we thank Matthew Luckie for insightful feedback on an earlier draft, our shepherd, Ellen Zegura, for help on the final version of the paper, and the anonymous reviewers for constructive critiques.

12. REFERENCES

[1] ALEXA. Alexa Top 1M Sites on 2012-10-02. http://www.alexa.com/topsites.

[2] ALLMAN, M. IPv6 DNS & Reachability Dataset. http://www.icir.org/mallman/data/ipv6-dns-data.tar.gz.

[3] APNIC PTY. LTD. APNIC's IPv4 Pool Usage, 2012. http://www.apnic.net.

[4] CAIDA. Archipelago Measurement Infrastructure. http://www.caida.org/projects/ark/.

[5] CHO, K., LUCKIE, M., AND HUFFAKER, B. Identifying IPv6 Network Problems in the Dual-stack World. In *Proceedings of the ACM SIGCOMM Workshop on Network Troubleshooting: Research, Theory and Operations Practice Meet Malfunctioning Reality* (2004).

[6] CLAFFY, K. Tracking IPv6 Evolution: Data We Have and Data We Need. *SIGCOMM Computer Communication Review* (2011).

[7] COLITTI, L., GUNDERSON, S. H., KLINE, E., AND REFICE, T. Evaluating IPv6 Adoption in the Internet. In *Proceedings of the 11th International Conference on Passive and Active Network Measurement*. 2010.

[8] CZYZ, J., ALLMAN, M., ZHANG, J., IEKEL-JOHNSON, S., OSTERWEIL, E., AND BAILEY, M. Measuring IPv6 Adoption. ICSI Technical Report TR-13-004, August 2013.

[9] DAVIES, J. Domain Name System Client Behavior in Windows Vista. http://technet.microsoft.com/en-us/library/bb727035.aspx, 2006.

[10] DHAMDHERE, A., LUCKIE, M., HUFFAKER, B., CLAFFY, K., ELMOKASHFI, A., AND ABEN, E. Measuring the Deployment of IPv6: Topology, Routing and Performance. In *Proceedings of the 12th ACM SIGCOMM Conference on Internet Measurement (IMC)* (2012).

[11] FEATHER, C. Network News Transfer Protocol (NNTP), 2006.

[12] FULLER, V., AND LI, T. RFC4632 - Classless Inter-domain Routing (CIDR): The Internet Address Assignment and Aggregation Plan, 2006.

[13] GIOTSAS, V., AND ZHOU, S. Detecting and Assessing the Hybrid IPv4/IPv6 AS Relationships. *SIGCOMM Computer Communication Review* (2011).

[14] GOOGLE. IPv6 Statistics. http://www.google.com/intl/en/ipv6/statistics/.

[15] GREGORI, E., IMPROTA, A., LENZINI, L., ROSSI, L., AND SANI, L. On the Incompleteness of the AS-level Graph: a Novel Methodology for BGP Route Collector Placement. In *Proceedings of the 12th ACM Conference on Internet Measurement (IMC)* (2012).

[16] GUÉRIN, R., AND HOSANAGAR, K. Fostering IPv6 Migration Through Network Quality Differentials. *SIGCOMM Computer Communication Review* (2010).

[17] GÜRSUN, G., RUCHANSKY, N., TERZI, E., AND CROVELLA, M. Inferring Visibility: Who's (Not) Talking to Whom? *SIGCOMM Computer Communications Review* (2012).

[18] HEI, Y., AND YAMAZAKI, K. Traffic Analysis and Worldwide Operation of Open 6to4 Relays for IPv6 Deployment. In *Proceedings of the International Symposium on Applications and the Internet* (2004).

[19] HUITEMA, S. RFC 4380: Teredo: Tunneling IPv6 over UDP through Network Address Translations (NATs), 2006.

[20] HUSTON, G. The Case for IPv6: Extinction, Evolution or Revolution? http://www.potaroo.net, 2006.

[21] IANA. IPv6 Addresses for the Root Servers. http://www.iana.org/reports/2008/root-aaaa-announcement.html, 2008.

[22] INTERNET SOCIETY. Internet Society - World IPv6 Day. http://www.worldipv6day.org/, 2011.

[23] KARPILOVSKY, E., GERBER, A., PEI, D., REXFORD, J., AND SHAIKH, A. Quantifying the Extent of IPv6 Deployment. In *Proceedings of the 10th International Conference on Passive and Active Network Measurement* (2009).

[24] LABOVITZ, C., IEKEL-JOHNSON, S., MCPHERSON, D., OBERHEIDE, J., AND JAHANIAN, F. Internet Inter-Domain Traffic. In *Proceedings of the 2007 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications (SIGCOMM)* (2010).

[25] LEBER, M. Global IPv6 Deployment Progress Report. http://bgp.he.net/ipv6-progress-report.cgi, 2014.

[26] MAHADEVAN, P., KRIOUKOV, D., FOMENKOV, M., DIMITROPOULOS, X., CLAFFY, K. C., AND VAHDAT, A. The Internet AS-level Topology: Three Data Sources and One Definitive Metric. *SIGCOMM Computer Communication Review* (2006).

[27] MARSAN, C. Biggest Mistake for IPv6: It's Not Backwards Compatible. http://www.networkworld.com/news/2009/032509-ipv6-mistake.html, 2009.

[28] MOCKAPETRIS, P., AND DUNLAP, K. J. Development of the Domain Name System. In *Proceedings of the 2007 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications (SIGCOMM)* (1988).

[29] NIKKHAH, M., GUÉRIN, R., LEE, Y., AND WOUNDY, R. Assessing IPv6 Through Web Access: A Measurement Study and Its Findings. In *Proceedings of the Seventh Conference on emerging Networking EXperiments and Technologies (CoNEXT)* (2011).

[30] OSTERWEIL, E., MCPHERSON, D., DIBENEDETTO, S., PAPADOPOULOS, C., AND MASSEY, D. Behavior of DNS' Top Talkers, a .com/.net View. In *Proceedings of the 13th International Conference on Passive and Active Network Measurement* (2012).

[31] PLONKA, D., AND BARFORD, P. Assessing Performance of Internet Services on IPv6. In *Proceedings of the IEEE Symposium on Computers and Communications (ISCC)* (2013).

[32] RIPE. RIPE Routing Information Service (RIS) Raw data Project. http://www.ripe.net/data-tools/stats/ris/ris-raw-data.

[33] RIPE NCC. Routing Information Service (RIS). http://www.ripe.net/ris/.

[34] SARRAR, N., MAIER, G., AGER, B., SOMMER, R., AND UHLIG, S. Investigating IPv6 Traffic - What Happened at the World IPv6 Day? In *Proceedings of the 13th International Conference on Passive and Active Network Measurement* (2012).

[35] SAVOLA, P. Observations of IPv6 Traffic on a 6to4 Relay. *SIGCOMM Computer Communication Review 35*, 1.

[36] SHEN, W., CHEN, Y., ZHANG, Q., CHEN, Y., DENG, B., LI, X., AND LV, G. Observations of IPv6 traffic. In *Proceedings of the International Colloquium on Computing, Communication, Control, and Management* (2009).

[37] SRISURESH, P., AND HOLDREGE, M. RFC2663 - IP Network Address Translator (NAT) Terminology and Considerations, 1999.

[38] UNIVERSITY OF OREGON. Route Views project. http://www.routeviews.org/.

[39] WANG, Y., YE, S., AND LI, X. Understanding Current IPv6 Performance: a Measurement Study. In *10th IEEE Symposium on Computers and Communications* (2005).

[40] WESSELS, D., LARSON, M., AND MANKIN, A. Analysis of Query Traffic to .com/.net Name Servers. http://www.apricot2013.net/__data/assets/pdf_file/0006/58884/130226-com-net-query-analysis-for-apricot-2013_1361840547.pdf, 2013. Slides of a talk given at APRICOT, Feb. 19–Mar. 1, Singapore.

[41] ZANDER, S., ANDREW, L. L., ARMITAGE, G., HUSTON, G., AND MICHAELSON, G. Mitigating Sampling Error When Measuring Internet Client IPv6 Capabilities. In *Proceedings of the 12th ACM Conference on Internet Measurement (IMC)* (2012).

[42] ZHANG, B., LIU, R., MASSEY, D., AND ZHANG, L. Collecting the Internet AS-level Topology. *SIGCOMM Computer Communication Review* (2005).

[43] ZHANG, G., QUOITIN, B., AND ZHOU, S. Phase Changes in the Evolution of the IPv4 and IPv6 AS-Level Internet Topologies. *Computer Communications 34*, 5 (Apr. 2011).

[44] ZHOU, X., AND VAN MIEGHEM, P. Hopcount and e2e Delay: IPv6 Versus IPv4. In *Proceedings of the 6th International Conference on Passive and Active Network Measurement* (2005).