Tracking the Deployment of IPv6: Topology, Routing and Performance

Siyuan Jia[ac*], Matthew Luckie[b], Bradley Huffaker[c], Ahmed Elmokashfi[d], Emile Aben[e], Kimberly Claffy[c], and Amogh Dhamdhere[c*]

Northeastern University, China[a], University of Waikato[b], CAIDA[c], Simula Research[d], RIPE NCC[e]

***Abstract*—We use historical BGP data and recent active measurements to analyze trends in the growth, structure, dynamics and performance of the evolving IPv6 Internet, and compare them to the evolution of IPv4. We find that the IPv6 network is maturing, albeit slowly, and notably IPv6 growth at the AS-level appears to have slowed after 2012. While most core Internet transit providers have deployed IPv6, edge networks are lagging. Early IPv6 network deployment was stronger in Europe and the Asia-Pacific region, than in North America. Current IPv6 network deployment still shows the same pattern. The IPv6 topology is characterized by a single dominant player – Hurricane Electric – which appears in a large fraction of IPv6 AS paths, and is more dominant in IPv6 than the most dominant player in IPv4. Routing dynamics in the IPv6 topology are largely similar to those in IPv4, and churn in both networks grows at the same rate as the underlying topologies. Our measurements suggest that performance over IPv6 paths is now largely comparable to (or better than) that over IPv4 paths.**

***Index Terms*—IPv6, BGP, topology, routing, performance.**

I. INTRODUCTION

The Internet operations, engineering and research communities are putting significant attention into a relatively new version of the Internet Protocol – IP version 6 (IPv6) [1] – designed to solve several architectural limitations of the existing IPv4 protocol. The most essential characteristic of IPv6 is that it provides orders of magnitude more address space than the world's foreseeable IP connectivity needs. IPv6 has become especially pertinent in the last few years because the global Internet address allocation architecture relies on the presence of a free pool of IP addresses to allocate to sites operating Internet infrastructure. The Internet Assigned Numbers Authority (IANA) exhausted its unallocated address pool in February 2011, and the Asia-Pacific region (represented by the APNIC RIR) followed suit in April 2011. The RIPE NCC, LACNIC and ARIN ran out of unallocated addresses in 2012, 2014, and 2015, respectively [2]. This exogenous pressure from IPv4 address scarcity has driven widespread adoption of IPv6 into modern operating systems and network equipment. Major network operators and content providers are deploying IPv6 on both a trial and production basis [3], and some governments are mandating IPv6 support [4], [5]. But there is little hard data about how mature the IPv6 network is in terms of composition, topology, routing, and performance.

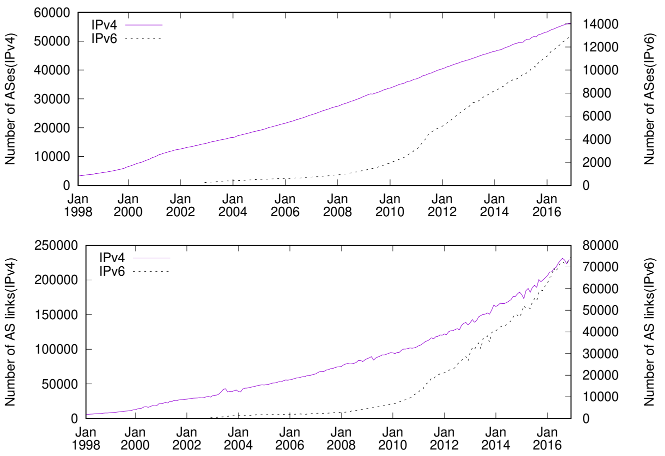


Fig. 1. Both AS nodes and links in IPv6 showed an initial exponential growth followed by a linear phase until the end of 2016. The IPv4 topology also displayed a similar transition from initial exponential growth to linear growth.

While IPv6 penetration is still small in comparison with IPv4, the IPv6 network topology showed signs of rapid initial growth. Figure 1 shows that IPv6 topology growth in fact displayed two distinct growth phases for both ASes and AS links, with a change in trajectory occurring around 2010-2011.

The trend shift shows that IPv6 growth has actually stalled in the last few years after the initial exponential growth. This is an interesting (and perhaps disturbing trend); in the previous version of this paper [6] which used data until 2012, we pointed to the exponential growth of IPv6 as strong evidence that IPv6 was maturing. Nevertheless, the continuing growth of IPv6 hints that it may finally have shifted from an experimental or "toy" network to production, so it is important to document various aspects of its growth, such as: Which network types and geographic regions contribute the most? Does the growing IPv6 network appear to converge toward the existing IPv4 network? How do routing dynamics in IPv6 compare to IPv4? How does performance over IPv6 paths compare with that over IPv4 paths? This paper provides an update to our earlier work which examined these questions in 2012. In this work, we expand our datasets until the end of 2016, and re-evaluate the extended dataset.

We use historical BGP archives and recent active measurements of the public IPv4 and IPv6 network infrastructures to analyze the state of maturity of IPv6 deployment along three

Corresponding author email addresses: jiasiyuan1116@163.com (Siyuan Jia), amogh@caida.org (Amogh Dhamdhere).

dimensions: topology, routing, and performance. Section II describes the data sources and supporting analysis techniques we use throughout the paper. We find that IPv6 deployment closely follows the footsteps of IPv4, as indicated by the growth trends in topology (AS nodes and links), business types, and geographical region. We find that the IPv6 network is maturing, as indicated by its increasing similarity to the public IPv4 Internet in size and composition (Section III), AS path congruity (Section IV), topological structure (Section V), and routing dynamics (Section VI). While core Internet transit providers have mostly deployed IPv6, edge networks are lagging behind. While all geographic regions showed early exponential growth in IPv6 adoption, early IPv6 deployment was stronger in Europe and the Asia-Pacific region than in North America. We find that in the 5 years since the previous version of our study, IPv6 growth has slowed down. The IPv6 AS topology now grows linearly in both ASes and AS links, a change from an initial exponential trajectory. This trajectory is consistent with that of IPv4, which also showed an initial exponential phase followed by linear growth until the present time. We find that Hurricane Electric (HE) is still the single prominent player in the IPv6 AS topology. Hurricane Electric currently appears in between 8% and 95% of IPv6 AS paths seen from different vantage points, and is more prominent in IPv6 than the most prominent player in IPv4. Further, when IPv4 and IPv6 AS path differ, HE is the network most often added to the IPv6 path. Routing dynamics in the IPv6 topology are largely similar to those in IPv4. While routing churn grows linearly in IPv4, it grew super-linearly in IPv6 during the exponential growth phase in the IPv6 topology. In other words, the trends in the growth of routing churn match those of underlying IPv4 and IPv6 AS topology growth. In terms of performance (Section VII), our performance measurements from 2017 show that IPv6 data-plane performance is comparable to (or better than) IPv4 performance. This is a significant change from our 2012 measurements, which showed that IPv6 was often worse (25%) than IPv4 performance.

II. DATASETS AND METHODS

We use a variety of data sources and analysis methods, which we summarize here, providing more detail in sections that use specific data. Our analysis of the IPv6 Internet's size, routing behavior, and structure (Sections III-V) relies on publicly available historical BGP routing data. Section VI uses BGP updates from the same public repositories to analyze routing dynamics of the IPv4 and IPv6 networks over time. We gather our own data using active measurements from four vantage points around the world, to compare and correlate IPv4 and IPv6 performance with other growth parameters. To compare the composition of the IPv4 and IPv6 graphs according to type of networks, we classify the business types of ASes using an algorithm similar to one presented in our previous work [7] and the business relationships of the links between them (*e.g.*, customer, provider, peer) using CAIDA's AS-rank algorithm [8].

BGP topology data: We collected historical BGP data from the two major public repositories at RouteViews [9] and RIPE [10]. We rely only on these two data sources because no other source of topological/routing data (routing registries, traceroutes, looking glass servers, etc.) provides historical information. Routeviews and RIPE started collecting IPv4 BGP data as early as 1998; the first IPv6 collector, however, became active in 2003. Consequently, our IPv4 data spans 18 years from 1998 to 2016, while the IPv6 data is from 2003 to 2016. The use of Routeviews/RIPE repositories of BGP data has been shown to inadequately expose the complete Internet topology [11]–[13] — although this data captures most ASes, it misses a significant fraction of peering and backup links at the edges of the Internet [13]–[15].

In this study, we use historical BGP data and construct a topology snapshot using all AS paths seen from all available Routeviews and RIPE vantage points in the first 5 days of each month. The time span is from Jan 1998 to Dec 2016, resulting in 228 snapshots of the IPv4 topology and 169 snapshots of the IPv6 topology.

BGP routing dynamics data: Our comparative analysis of routing dynamics of the IPv4 and IPv6 infrastructures is based on BGP updates collected by the Routeviews project. Routeviews collectors run BGP sessions with routers (or monitors) in many networks. Each monitor sends a BGP update to the collector every time there is a change in the preferred path from the monitor to a destination prefix. We use update traces from two Routeviews collectors: Routeviews6 for IPv6 data and Oregon-IX for IPv4 data. The IPv4 updates span the period from 1 Jan 2003 to 31 Dec 2016, while the IPv6 updates span 7 May 2003 through 31 Dec 2016. We use monitors from five networks that contributed both IPv4 and IPv6 routing data throughout the study period: AT&T, Hurricane Electric (HE), NTT-America, and Tinet, and IIJ. AT&T's IPv4 monitor was unavailable for three months in 2003, and its IPv6 monitor was unavailable between May 2005 and May 2007. Tinet's IPv6 monitor was unavailable between June 2008 and June 2010. If the multi-hop BGP session between a monitor and the collector is broken and re-established (session reset), the monitor re-announces all its known paths, producing large bursts of updates. This is a local artifact of the measurement infrastructure, and does not represent genuine routing dynamics. We use the method developed by Zhang et al. [16] to identify and remove updates caused by session resets.

AS relationships: We use CAIDA's AS-rank AS relationship classification algorithm [8] to infer the business relationship associated with each inter-AS link. For each snapshot, we apply this algorithm to the set of IPv4 AS paths. The AS-rank algorithm classifies AS links into two types: customer-provider or settlement-free peer.

AS classification: We use a machine learning approach similar to a method from our previous work [7] to classify ASes according to their business type. Our method consists of using labeled AS classification data to train a machine-learning classifier to classify ASes according to their business type. We first use a ground-truth dataset from PeeringDB [17], and split it into two parts to create labeled training and validation sets. PeeringDB is the largest source of self-reported data about the properties of ASes. From PeeringDB, we extract the self-reported business type of each AS, which is one of “Cable/DSL/ISP”, “NSP” (Network Service Provider), “Content”, “Education/Research”, “Enterprise” and “Non-profit”. We combine the “Cable/DSL/ISP” and “NSP” classes into a single class “Transit/Access”. We ignore the “Non-profit” category for the purposes of this classification. The labeled ground-truth data thus consists of three classes: “Transit/Access (TA)”,“Content Provider (CP)” and “Enterprise Customer (EC)”. As PeeringDB under-represents the “EC” category, we manually assemble a set of 500 networks which we determine to be enterprise customers based on their WHOIS records and webpages, and add this set to the labeled classification data. We then train a machine-learning classifier using the following features for each AS: 1) Customer, provider and peer degrees: We obtain the number of customers, providers and peers (at the AS-level) using CAIDA’s AS-rank data [18]; 2) Size of customer cone in terms of number of ASes: We obtain the size of an AS’ customer cone using CAIDA’s AS-rank data [18]; 3) Size of the IPv4 address space advertised by that AS. We obtain this quantity using BGP routing tables collected from Routeviews; 4) Number of domains from the Alexa top 1 million list [19] hosted by the AS. We obtain the list of top 1 million websites from Alexa, perform DNS lookups on each domain (at CAIDA) and map each returned IP address to the corresponding ASN using longest-prefix matching using a routing table from Routeviews. We then count the number of domains hosted by each AS; 5) Fraction of an AS’s advertised space that is seen as active in the UCSD Network Telescope [20]. This AS classification method has a accuracy (positive predictive value) of 70% [21]. **Performance data:** Similar to the method employed by Nikkhah et al. [22], we measure the average time to fetch a page from webservers with IPv4 and IPv6 addresses that have the same origin AS number in BGP. We use the Alexa list of the one-million most popular websites in the Internet, testing up to three webservers for each origin AS. From each webserver, we seek to download a page that is at least 100KB. If the web site’s root page is smaller than that, we fetch the smallest object embedded in that page that is at least 100KB. While a threshold of 100 KB may not always be sufficient to get out of slow start, it seeks to balance the tradeoff between finding a large number of web objects to download, and ensuring that those objects are sufficiently large. We fetch each page three times from each webserver, alternating IPv4 and IPv6 transport sequentially. Each measurement begins approximately five seconds after the previous one completes to avoid competing measurements but also to minimize the chance of network topology changes mid-measurement. We also measure the forward AS-level IPv4 and IPv6 paths using traceroute with TCP probes immediately after the sequence of performance measurements completes. We collected this data from four vantage points in October 2017: An Internet Exchange in Canada, an educational network in Finland, a research network in Ireland, and a commercial network in the US. We sanitized our measurements as in [22]: We excluded measurements where the standard error of the mean download time (for either IPv4 or IPv6) was greater than 10% (at the 95% significance level), or the object sizes in IPv4 and IPv6 were not within 1% of each other. This filtering left us with 705 dual-stack ASes represented in our dataset, consisting of 105 ECs, 361 TAs and 230 CPs according to the previously described classification. We used scamper’s tbit and traceroute implementations [23]; the former includes a test that fetches a page, negotiating TCP SACK and TCP timestamps, and records all packets sent and received during the test, which allows us to further examine the packet traces to infer why performance may differ.

III. GROWTH TRENDS BY BUSINESS TYPE AND GEOGRAPHIC REGION

While overall growth rates indicate that IPv6 deployment is growing, growth differs by (business) type of network and geographic region. Since IPv6 provides essentially the same functionality as IPv4, we hypothesize that as IPv6 matures, the distribution of business types in IPv6 should resemble that in IPv4. Geographic coverage of IPv6 may not exhibit the same convergence with IPv4 given the pre-existing allocation of IPv4 address space around the world and various levels of pressure by different national governments to promote IPv6.

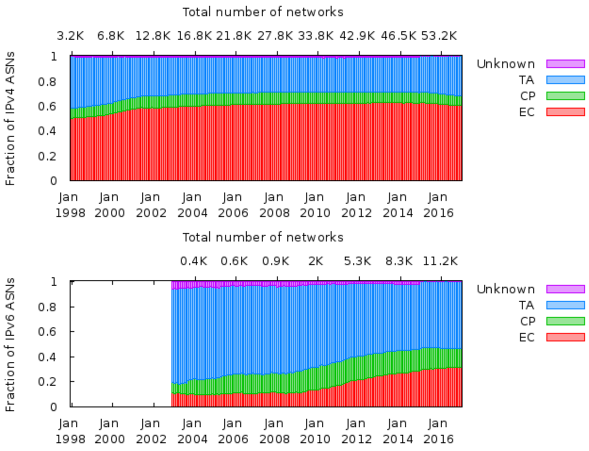


Fig. 2. As IPv6 matures the fraction of EC ASes has grown from 11.65% to 31.93% of the IPv6 graph, while IPv4 has seen little change, with ECs currently at 60.40%. This suggests that IPv6 deployment is strong in the core of the network, while it lags at the edge.

A. *Growth Trends by Business Type*

Figure 2 shows the fraction of networks over time from each of the three business types (and an “unknown” type) mentioned in Section II, for the IPv4 (top panel) and IPv6 (bottom panel) topologies. Above each panel, we show the total number of ASes in the IPv4 and IPv6 graphs over which the fractions are computed. In 1998, 50.46% of IPv4 networks were of type EC, 7.85% of type CP and 41.65% were of type TA. At the end of 2016, 60.40% of ASes were EC, 31.85% of ASes were TA and CP comprised most of the remaining 8%. In 2003, 75.1% of IPv6 networks were of type TA, but this fraction had reduced year after year, and in Dec 2016 the fraction was 53.51%. The fraction of IPv6 networks of type EC has increased steadily from 11.65% in 2003 to 31.93% at the end of 2016. The fraction of type CP has climbed from

13.25% in 2003 to 14.55% in Dec 2016. *The large fraction of TA in the IPv6 topology suggests that IPv6 deployment has primarily occurred at the core of the network, driven by transit and access providers, not by CPs or ECs.*

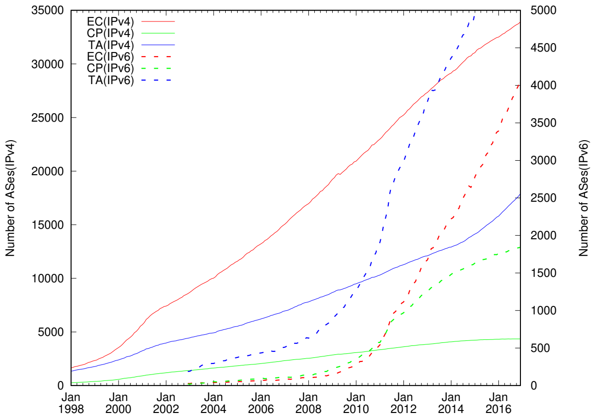


Fig. 3. Growth of the number of ASes in the IPv4 and IPv6 AS graphs over time, classified by business type. The growth trend of all business types changed from exponential to linear in both the IPv4 and IPv6 topologies (as of 2010-2011).

To further explore the evolution of business types in IPv6, we measure growth trends for each AS type in the IPv4 and IPv6 graphs in Figure 3. We find that ECs, CPs and TAs in IPv4 all showed an initial exponential growth followed by a shift to linear growth in Oct 2002, Aug 2001 and May 2003, respectively. The IPv6 graph has evolved similarly. For IPv6 ECs, TAs and CPs, we observe that an exponential growth phase from 2003 (when data archiving began) until 2010-2011, followed by linear growth gives the best fit with the data. The exponents for ECs, TAs and CPs in the exponential growth phase are 0.0339, 0.02552 and 0.037, respectively. We also find that the number of ECs exceeded the number of CPs in IPv6 in June 2011. We also measure the growth rate of each business type in the IPv4 and IPv6 graphs. In IPv4, the growth rate of business type is divided into two parts. From 1998 to 2002, the IPv4 network showed a high fluctuation and the average growth rate is 0.03. After 2002, the growth rate was slow down and the average growth rate is down to 0.009. And this inflection time correspond to the exponential growth following with linear growth in Figure 3. In IPv6, the growth rates of different AS types shows a clear change point around June 2011 and the average growth rate is 0.033, 0.020 separately. Before 2011, the growth rate of each business type were very small (TAs between -9 and 50 ASes/month; ECs between -5 and 15 ASes/month; CPs between -5 and 15 ASes/month). After 2010, the growth rate of each business type increased significantly (TAs between -27 and 204 ASes/month; ECs between -18 and 82 ASes/month; CPs between 0 and 62 ASes/month). The growth rate of TAs, ECs and CPs in the IPv6 reached a peak of 204 ASes/month, 78 ASes/month, and 62 ASes/month, respectively in June 2011 which also coincident with World IPv6 Day [24]. *In the previous version of this paper [6], we had observed that all business types in IPv6 were growing exponentially as of 2012. The updated data reveals that the growth of all business types in IPv6 has slowed down since then, settling into a slower linear growth. The growth trend of IPv6 now more closely resembles that in IPv4.*

B. Growth Trends by Geographical Region

Figure 4 shows the number of ASes in different geographical regions over time, according to the RIR WHOIS mappings described in Section II. We put the two smallest registries (LACNIC and AFRINIC) in APPENDIX (Figure 20), as they have so few ASes compared to the three large registries (ARIN, RIPE and APNIC) that they are barely visible in the graph. The graph shows that for IPv4, the growth rate of RIPE-registered ASes has exceeded that of ARIN-registered ASes for the last decade (though both ARIN and RIPE showed linear growth in this period), and as of 2009 the RIPE region has more ASes than the ARIN region, a big difference from the early days of IPv4.

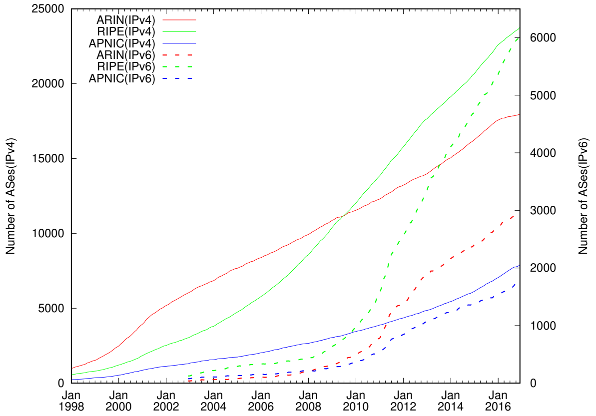


Fig. 4. Regional growth in IPv4 and IPv6 ASes. The RIPE region has the most ASes in both IPv4 and IPv6.

On the other hand, the growth trend in IPv6 for each of the ARIN, RIPE and APNIC registries shows two distinct periods since 2003 - an initial exponential phase (with exponents 0.038, 0.028 and 0.022, respectively) followed by an a linear phase until Dec 2016. For ARIN and APNIC, the changes from exponential to linear happened in May 2011, while for RIPE it was at the start of Jan 2011. Unlike IPv4, however, the RIPE region has always had more ASes in IPv6 than ARIN. After 2011, the growth rate in RIPE-registered, ARIN-registered and APNIC-registered slowed down. The regional growth of IPv4 and IPv6 ASes classified by business type are provided in the APPENDIX (Figure 21 and Figure 22).

IV. EVOLVING STRUCTURE OF IPV4 AND IPV6 TOPOLOGIES : AS PATHS

Similar to our belief that the composition of a maturing IPv6 topology should look more like the IPv4 topology, we also expect a convergence to occur between the best AS path between a given pair of ASes in IPv4 and IPv6. Another reason to compare IPv4 and IPv6 AS path congruity is its correlation with performance. In Section VII we show that *IPv6 data plane performance is largely comparable to IPv4,* but performance in both protocols is more similar when the AS paths are the same. Improved congruity between IPv4 and IPv6 paths seem to lead to more similar performance between IPv4 and IPv6 performance, which is likely to further promote IPv6 deployment. For the analysis presented in this section, we use a set of BGP vantage points to present our results. We have chosen VPs that are located in a diverse set of networks in terms of size and position in the Internet hierarchy, also ensuring that these VPs could provide us a longitudinal view from 2003 to present. As a result, we cannot conduct this analysis for a large set of VPs. Hence it should be noted that the results presented here may not generalize to all ASes on the Internet.

To explore trends in congruity between IPv4 and IPv6 paths, we first calculate the fraction of AS paths from a given vantage point (VP) toward dual-stacked origin ASes (*i.e.*, ASes that advertise both IPv4 and IPv6 prefixes) that are identical in IPv4 and IPv6. Let $O = \{o_1, o_2, ..., o_N\}, N \geq 1$ be the set of dual-stacked origin ASes observed from a given VP. Let P_i^4 and P_i^6 be the set of AS paths observed for one origin AS o_i $(1 \leq i \leq N)$ in IPv4 and IPv6, respectively. If $\exists\ P_i^4 \cap P_i^6 \neq \emptyset$, we record $|o_i| = 1$. Otherwise, we record $|o_i| = 0$. So, the *fraction of identical AS paths* ($FIAP$) between IPv4 and IPv6 is:

$$FIAP = \frac{\sum_{i=1}^{N} o_i}{N} \tag{1}$$

If there are multiple IPv4 or IPv6 AS paths available between a given VP and an origin AS, we report the VP as having an identical AS path if any of the paths are the same. If they differ, we dissect the differences, in terms of which ASes are added and removed from those paths. This analysis also reveals the presence of dominant players in the IPv6 topology. So we use seven vantage points listed in Table I which have provided BGP data to Routeviews and RIS since 2003. For each topology snapshot, we use all AS paths exported by these monitors in the first 5 days of each month (as described in Section II). We remove all prepending from AS paths, and discard paths with AS sets or loops; this filtering rejects 0.1% of AS paths.

Identical AS paths in IPv4 and IPv6: Figure 5 plots the fraction of dual-stack paths that are identical in IPv4 and IPv6 from different vantage point over time. According to this metric, IPv6 paths have matured significantly over the last decade. In January 2004, 10-20% of paths were the same for IPv4 and IPv6; In Dec 2016, more than a decade later, 40-75% of paths are the same. There are, however, significant differences across monitors. For five of the 7 VPs (NTT, Tinet, HE and IIJ, AT&T), the fraction of identical paths is higher, between 64 and 75%. For NL-BIT and ACOnet, however, the fraction of identical paths is lower, 53% and 43% respectively.

An interesting trend in this data is the rise in prominence of Hurricane Electric. In April 2007 only 5% of its dual-stacked paths were identical, but in 2012 it occupied the top position, with just over 50% of dual-stacked paths identical from Hurricane's perspective. Since 2012, however, the growth of path congruity appers to have slowed down for HE; consequently, NTT had a larger fraction of identical paths than HE at the end of 2016.

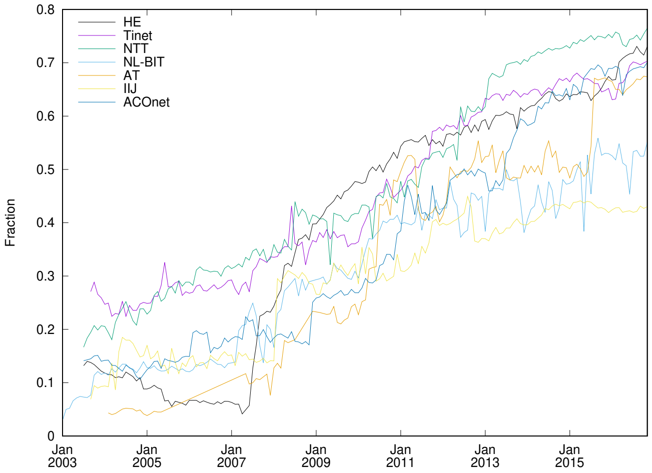


Fig. 5. Fraction of dual-stacked origin ASes reachable over an identical AS-level path in both IPv4 and IPv6. At the end of 2016, more than 62% of the AS-level paths used to reach an origin AS are the same in both protocols. The fraction of identical paths from NTT has exceeded that of HE, which held the top position from 2008 to 2012.

Different ASes in IPv4 and IPv6 AS paths: Since between 40% and 75% of the AS paths from different vantage points to dual-stacked origin ASes are the same, the next question is: how do the paths differ? We compute the AS edits required to make IPv4 paths identical to IPv6 paths — specifically, which ASes are most often added and removed from AS paths that differ. Considering the IPv4 and IPv6 paths from a monitor to the same origin AS in a given snapshot, we define an AS as "added" to the IPv6 path, if the AS is present in the IPv6 path but not present in IPv4 path. Similarly we define an AS as "removed" from the IPv4 path if that AS is present in the IPv4 path but not in the IPv6 path. For both added ASes and removed ASes, we compute the fraction of paths seen in that snapshot in which those ASes were added or removed. In order to ensure that the results were not affected due to short-lived phenomenon in a monthly snapshot, we repeated this procedure for 12 monthly snapshots from 2016, and computed the average fraction of paths in which ASes were added and removed across these 12 snapshots. Furthermore, it should be noted that the results presented here could depend on the position of VPs.

Figure 6 shows the average fraction (over all snapshots in 2016) of AS paths from which an AS was added or removed, for all ASes that were in the top-10 added or removed ASes for at least one snapshot in 2016. Note that the top-10 ASes differ across snapshots (an AS may be among the top-10 in one month, but falls off that list in the next month), so the number of ASes we plot is larger than 10. The upper panel of the figure shows the average fraction of paths in which an AS was added, and the x-axis shows all ASes that were in the top-10 in any of the 12 snapshots (a total of 31 ASes). Hurricane Electric (AS 6939) was added most frequently, with an average fraction of over 45%; other ASes, on the other hand, were added to less than 10% of paths. Hurricane Electric was thus the AS most frequently and consistently added to IPv6 paths across our seven vantage points. The bottom panel of Figure 6 shows the

TABLE I
BGP VANTAGE POINTS (VP) PROVIDING BOTH IPV4 AND IPV6 ROUTING DATA SINCE 2003.

Peer	ASN	Name	Type	BGP source	when
ACOnet	1853	Austrian Academic Computer Network	TA	RIS RRC 05	Oct 2003
IIJ	2497	Internet Initiative Japan	TA	Routeviews 2/6	Jul 2003
NTT	2914	NTT Global IP Network	TA	Routeviews 2/6	Jul 2003
Tinet	3257	Tiscali International Network	TA	Routeviews 2/6	Oct 2003
HE	6939	Hurricane Electric	TA	Routeviews 2/6	Jul 2003
AT&T	7018	AT&T Services	TA	Routeviews 2/6	Apr 2003
BIT	12859	BIT BV	TA	RIS RRC 03	Jan 2003

average fraction of AS paths from which an AS was removed in IPv4 paths, for the top-10 removed ASes during any month. Level 3 (AS 3356) was the AS most often removed from IPv4 paths (32%), while NTT (AS 2914) was close behind at 28%. Therefore, from the perspective of BGP vantage points chosen in this analysis, Hurricane Electric and Level3 were the most frequently added/removed ASes in IPv6/Ipv4 paths.

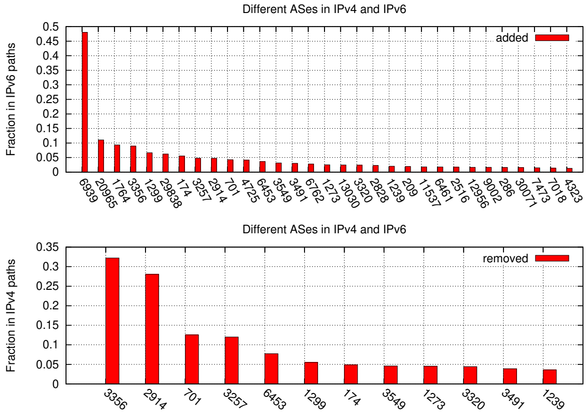


Fig. 6. The average fraction of AS paths to which ASes were added, for the top-10 added ASes in any snapshot in 2016 (top panel), and fraction of paths from which ASes were removed (bottom panel). Hurricane Electric was the most frequently and consistently added AS in IPv6 paths across vantage points. Level 3 and NTT were the most frequently removed ASes.

ASes most frequently seen in AS paths: Next, we examine the AS paths from all BGP vantage points (VPs) that provide a full table to Routeviews and RIPE collectors in Dec 2016 to determine the relative *prominence* of ASes in the IPv4 and IPv6 topologies. We define the *prominence* of an AS X to a VP as the fraction of origin ASes that are reached through it. As of Dec 2016, 125 of the VPs are dual-stacked and they provide both IPv4 and IPv6 full BGP views. For a VP, let N_4 be the number of origin ASes observed in paths from this VP in IPv4. Let N_4^{3356} be the number of origin ASes reached by the VP via Level3 in IPv4. We denote f_4^{3356} as the fraction of origin ASes reached by the VP via AS3356 in IPv4: $f_4^{3356} = N_4^{3356}/N_4$. Similarly, f_4^{6939} represents the fraction of origin ASes reached by the VP via AS6939 in IPv4. Finally, we also compute these fractions for each VP in IPv4, i.e., f_6^{3356} and f_6^{6939}. In Figure 7, the X axis represents the index of each of the 125 VPs. The Y axis shows the fraction of origin ASes reached by the VP via an AS (either Hurricane Electric or Level3) in IPv4 and IPv6, i.e., we show each of the 4 fractions computed for each VP. While the AS that appears most often depends on the VP in question, Figure 7 shows that Hurricane Electric appears in between 7% and 98% of IPv6 AS paths (with an average of 55.28%) depending on the VP. Contrast this with the importance of Hurricane Electric in the IPv4 topology, where it appears in fewer than 20% of AS paths for 90% of the VPs. Level3 (AS3356), the largest player in the IPv4 space in terms of this metric, appears in between 5% and 86% of IPv4 AS paths (with an average of 28.48%) depending on the VP. This data suggests that *Hurricane Electric is more prominent in the IPv6 graph than the most prominent player in the IPv4 graph.*

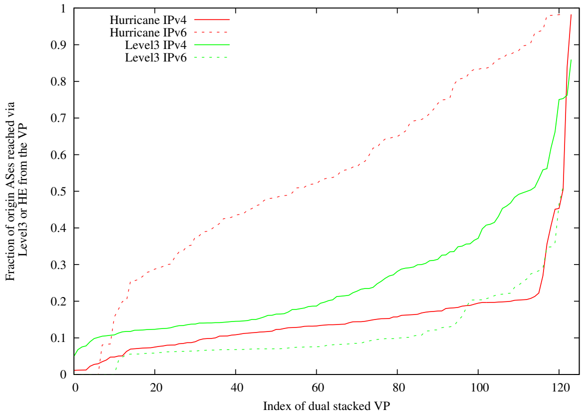


Fig. 7. Fraction of origin ASes reached via an AS (Hurricane Electric, Level3) for each BGP vantage point in Dec 2016. Hurricane is relatively more prominent in the IPv6 topology than Level3 is in the IPv4 topology.

AS path lengths in IPv4 and IPv6: Even though the IPv4 AS graph continues to grow in the number of ASes (linearly, after initial exponential growth until 2002), the average AS path length as measured from Routeviews/RIPE vantage points is *almost constant around 4 AS hops* since January 1998 [7]. We emphasize that this result is based on ASes that provide data to Routeviews and RIPE collectors, and does not necessarily reflect the average AS path length that an arbitrary AS sees. Figure 8 shows the average path length in the IPv4 and IPv6 topologies over time. The average AS path length for IPv6 showed an initial decreasing trend, with a sharp decrease

We define a VP as having a full table if it has BGP paths to at least 50,000 IPv4 ASes and 11,000 IPv6 ASes as of Dec 2016.

This metric is related to *betweenness centrality*, but only uses paths observed from a single VP.

around 2008. After 2008 the average IPv6 path length has been mostly stable. The average IPv4 path length has been fairly stable since 2004.

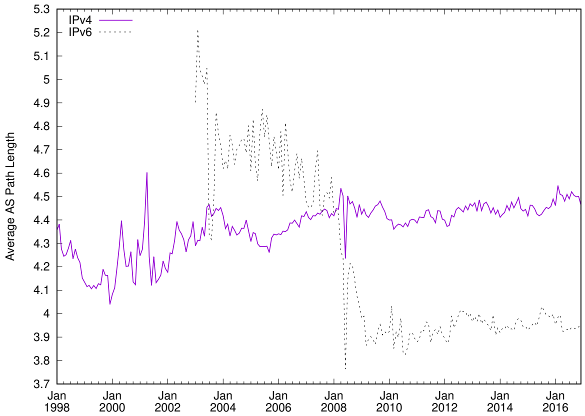


Fig. 8. Average IPv4 AS path length is almost constant, while in IPv6 it decreased until 2008 and has been stable since then.

What caused the sudden decrease in IPv6 AS path length around 2008? Figure 9 shows the average AS path length seen from the perspective of Hurricane Electric (9(a)), and from the other vantage points (9(b)). The average path lengths from vantage points other than HE are similar, hence we group them together. The plot also shows the number of AS paths of length 2 (origin AS is directly connected to the VP), of length $\leq$ 3, and so on. Our main observation is that the average IPv6 AS path length from HE showed a sharp decrease around 2008, while the path length measured from other VPs has been more constant. The fraction of ASes directly connected to Hurricane Electric in IPv6 increased sharply from 5% in 2007 to a peak of 37% in 2010, followed by a slow decrease until the present time. This is perhaps as a result of Hurricane's open peering policy in IPv6 [25]. Since 2003, other transit providers have observed the fraction of directly connected dual-stack ASes decrease (indicated by the curve labeled "==2" in Figure 9(b)), further confirming the rising dominance of HE in the IPv6 topology. We conclude that the sharp decrease seen in the average IPv6 AS path around 2008 is due to this observed dominance of HE in the IPv6 topology. We recommend caution in analyzing graph properties of the IPv6 AS topology; due to its relatively small size, the presence of even a few important ASes such as Hurricane Electric can significantly affect overall graph properties.

V. EVOLVING STRUCTURE OF IPV4 AND IPV6 TOPOLOGIES : AS GRAPHS

We next directly compare IPv4 and IPv6 topologies over time. Again we hypothesize that as the IPv6 network matures, its topological structure should grow more congruent with that of IPv4, i.e., an increasing fraction of ASes and AS links will be common to both topologies, the most highly connected ASes should grow to be more similar in both topologies, and upstream IPv4 and IPv6 providers for the same edge AS should eventually converge.

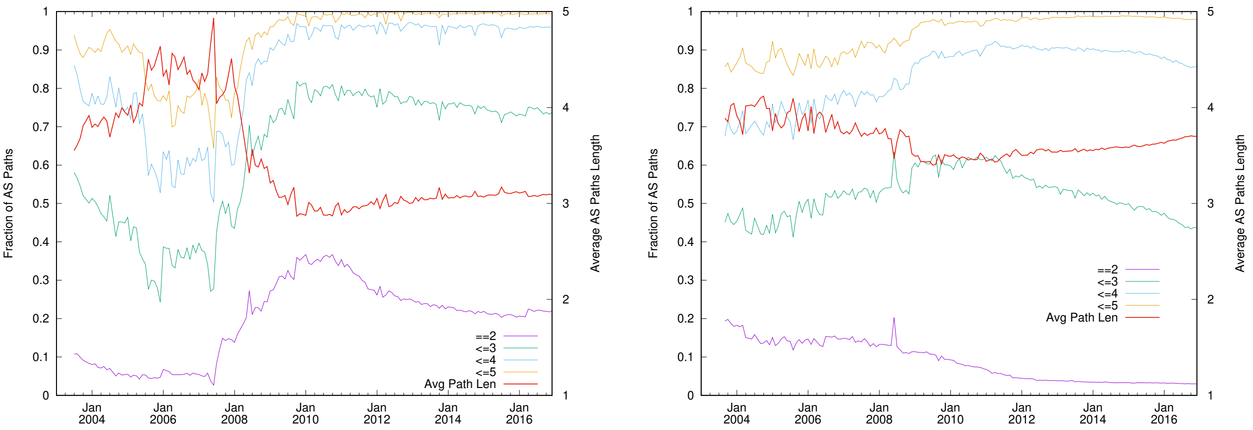
A. *Common ASes and AS links in IPv4 and IPv6 graphs*

For each topology snapshot, we find the set of ASes that are present in either the IPv4 or the IPv6 AS topology, which we call the combined topology. In each snapshot, more than 99% of ASes and AS links in the combined topology were present in the IPv4 topology, *i.e.*, *the number of ASes that are unique to the IPv6 topology is negligibly small*. Consequently, we focus most of our analysis on the set of ASes from the combined topology that were present in the IPv6 topology.

Common ASes present in the IPv6 topology: Figure 10 shows the fraction of ASes from the combined topology that are present in the IPv6 topology. We measure these fractions for all ASes, and further classify ASes according to business type. We find that the fraction of ASes from the combined topology that are seen in the IPv6 topology varies widely depending on business type. We also find some interesting trends over time. Initially, TA networks had the largest fraction of ASes that were present in the IPv6 topology. Since 2011, however, CP networks have the highest fraction of common ASes, and this fraction was approximately 43% at the end of 2016. Curiously, this transition happened around the same time that the growth rates of CP and TA AS types in IPv6 changed from exponential to linear. At the end of 2016, fewer than 12% of ECs were seen in the IPv6 topology. Since the combined AS topology is dominated by ECs, the overall fraction of ASes from the combined topology that are seen in the IPv6 topology is similarly low (less than 23% at the end of 2016), which confirms our earlier observation that IPv6 adoption is faster in the core of the network while the edge (ECs) has been slow to deploy IPv6.

We also measured the fraction of ASes from the combined topology that are present in the IPv6 topology, separately for each geographic region, as shown in Figure 11. Prior to 2014, the APNIC region had the largest fraction of ASes that were in the IPv6 topology, higher than both RIPE NCC and ARIN. However, RIPE NCC overtook APNIC in Nov 2014. We find that these fractions show similar growth trends for all geographical regions as the overall IPv6 topology — an exponential growth phase peaking around 2011, followed by slower linear growth. Currently, the RIPE region has the largest fraction of ASes in the IPv6 topology; 23% of ASes from the combined topology are present in IPv6 for RIPE NCC.

ASes unique to the IPv6 topology: We briefly comment on the small set of ASes that were present only in the IPv6 topology. In our latest topology snapshot from Dec 2016, 334 ASes were only in the IPv6 topology. Of these, 149 ASes (66 ECs, 22 TAs, 5 CPs and 27 Unknowns) were in the IPv4 topology in some previous snapshot. Inspection of the AS names and descriptions of the other 185 ASes (as they appear in the RIR whois database) reveals that 27 can be trivially matched with ASes in the IPv4 topology that have similar names and descriptions. This overlap hints at organizations using separate ASes to provide IPv4 and IPv6 connectivity. Furthermore, we found that 13 ASes unique to the IPv6 topology were administered by universities that used IPv4 address space announced from the respective national research and education networks ASes. This shows that organizational



(a) Hurricane Electric (b) Combined Tinet, NTT, IIJ IPv6.

Fig. 9. Average AS Path lengths to dual-stacked origin networks over time from different vantage points, and the fraction of paths of length 2 (directly connected), length <= 3, and so on. In October 2010, HE was directly connected to 37% of dual-stacked origin ASes in IPv6. Since 2003, other transit providers have observed the fraction of directly connected dual-stack ASes decrease.

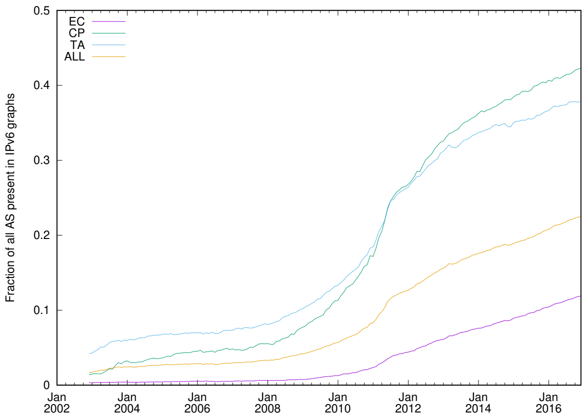


Fig. 10. Fraction of ASes from the combined (IPv4+IPv6) graph that are present in the IPv6 graph, classified according to business type.

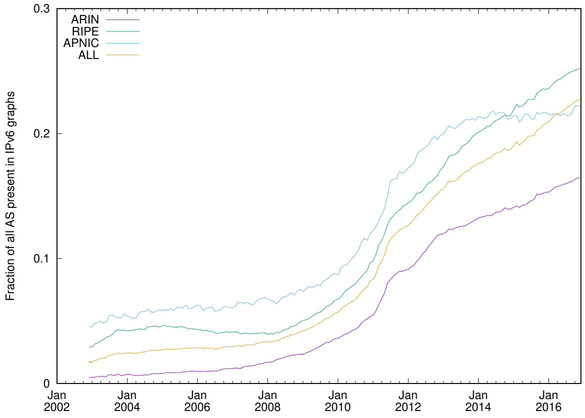


Fig. 11. Fraction of ASes from the combined (IPv4+IPv6) graph that are present in the IPv6 graph, classified according to geographical region.

boundaries of the entities that manage ASes in the IPv4 and IPv6 topology do not always align.

Common top ASes: We measure the fraction of the top-K ASes (in terms of AS degree) from the IPv4 topology that are also top among the top-K ASes in the IPv6 topology. As the IPv6 network matures, we expect that the top ASes from the IPv4 topology will also appear as the top ASes in the IPv6 topology. Figure 12 shows the fraction of the top-K ASes from the IPv4 topology also among the top-K ASes in the IPv6 topology, for K=10, 50 and 100. The fractions have increased from less than 20% in 2003 to more than 60% for K=50,100 100 and 80% for K=10 in 2016. Until 2011, however, the top-K fraction for K=10 was almost similar with that for K=50 and K=100. After 2011, the top-K fraction for K=10 has been between 80% to 100%, while the fraction for K=50 and K=100 were from 50% to 70%, indicating that the largest ASes in the IPv4 topology were present in IPv6, and were among the largest ASes in the IPv6 topology.

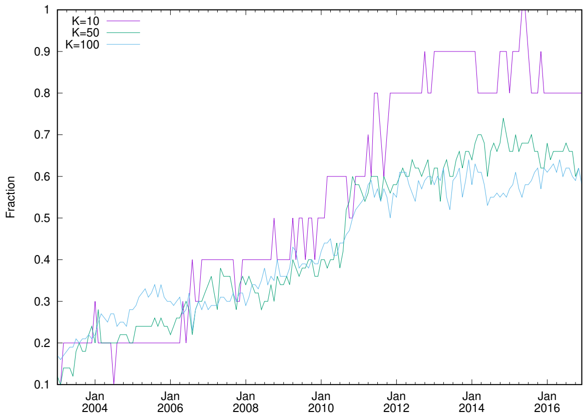


Fig. 12. Fraction of ASes from the top-K ASes in the IPv4 graph that are also in the top-K ASes in the IPv6 graph. The fractions increase for different K values.

Common AS links: Finally, we are interested in the common AS links between the IPv4 and IPv6 topologies. As mentioned in Section II, our BGP vantage points are likely to miss AS links, particularly peering links lower in the hierarchy than our vantage points. We are, however, interested in the fraction of links from the IPv4 topology which were also seen in the

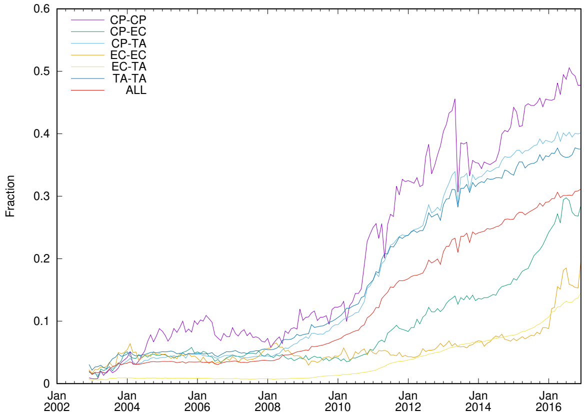


Fig. 13. Fraction of AS links from combined (IPv4 + IPv6) graph that are present in the IPv6 graph, classified according to business type of link endpoints.

IPv6 topology. Visibility issues should affect both the IPv4 and IPv6 graphs similarly, and hence our analysis should not be impacted by missing peering links in the measured IPv4 and IPv6 topologies.

Figure 13 shows the fraction of AS links from the combined topology that also appear in the IPv6 topology over time. We compute this fraction for all AS links and also according to the business types of the endpoints. The fraction of AS links seen in the IPv6 topology was rather low, at around 30% in Dec 2016. The fraction of common AS links varies significantly according to the business type of the endpoints, however. Links involving ECs are the least represented in the IPv6 graph, while larger fraction of links involving CPs and TAs are seen in the IPv6 graph. This is again consistent with our previous finding that the pace of IPv6 adoption is higher in the core of the network but lags at the edge (represented by ECs).

VI. EVOLVING DYNAMICS OF IPV4 AND IPV6 INFRASTRUCTURE

Continuing to explore our hypothesis that a maturing IPv6 network should look more like the IPv4 network, we compare the evolution of routing dynamics in IPv4 and IPv6. In particular, we focus on the evolution of update churn, correlation between the update churn seen from different vantage points, path exploration, and convergence times in IPv4 and IPv6. We focus on these metrics for the following reasons. First, we hypothesize that both IPv4 and IPv6 should show a similar relation between update churn and the size of the underlying topology. Second, due to business relationships and dense interconnection among ASes, churn becomes localized, and each vantage point does not see the same set of routing events. Consequently, correlation between update churn seen at different vantage points can serve as a measure of the maturity of the underlying network and business relationships. Finally, previous work has shown that end-to-end delays and loss rates are significantly higher during routing events [26]. It is thus useful to compare the extent of path exploration and routing convergence times during routing events. If these metrics are significantly worse in IPv6 as compared to IPv4, then it could deter the adoption of IPv6. Similarly to the analysis in Section IV, we use a subset of BGP vantage points to present our results on routing dynamics. We have chosen VPs that are located in a diverse set of networks in terms of size and position in the Internet hierarchy, also ensuring that these VPs could provide us a longitudinal historical view of routing updates. Hence it should be noted that the results presented here may not generalize to all ASes on the Internet.

A. *Churn as a function of topology size and vantage point*

Interdomain routing scalability has been a topic of major concern in recent times [27], [28] for two reasons – increasing routing table size, and increasing rate of BGP updates (churn). The latter can be a more serious concern, because failing to process updates in a timely manner can trigger a wide-scale instability and result in traffic blackholing. Some of these concerns were put to rest by observations that churn in the IPv4 topology grows slowly [29], [30], and at the same rate as the underlying topology. More recently, however, Huston [31] compared IPv4 and IPv6 BGP update time series and concluded that while IPv4 churn has grown slowly (linear), IPv6 churn has been increasing exponentially. This qualitative difference between the evolution of update churn in IPv4 and IPv6 raised speculation on whether routing dynamics in IPv6 are fundamentally different from those in IPv4. In order to investigate these differences, we next compare the evolution of BGP churn in IPv4 and IPv6. We define churn as the rate of BGP updates received from a vantage point (*e.g.*, updates per day). This definition of churn is consistent with previous related work in the area.

Churn as a function of topology size: To understand how churn has evolved with respect to network size, we track the growth in the number of updates, normalized by the size of the underlying AS topology. To calculate this metric, we bin the total number of updates per day into three-month windows, find the median daily churn (using the average daily churn gives similar results) for each window, and divide it by average number of ASes in the topology during that time window. We define a set C to represent the total number of daily updates over three months as : $C = \{c_1, c_2, ..., c_N\}$. Then we sort it according to the order from small to large as: $C_{sort} = \{c_{(1)}, c_{(2)}, ..., c_{(N)}\}$. Next, the median daily churn in three-month window is

$$Mid_{(X)} = \begin{cases} X_{(N+1)/2} & \text{if } N \text{ is odd number} \\ \frac{X_{(N/2)}+X_{(N+1)/2}}{2} & \text{if } N \text{ is even number} \end{cases} \quad (2)$$

Last, we normalized the median daily churn by the average number of ASes in the topology during that time window which mentioned in Section II. Let N_1, N_2, N_3 be the number of ASes in the topology in these three months. Therefore, the churn growth in relation to toplogy size is:

$$\frac{Mid_{(C)}}{(N_1 + N_2 + N_3)/3} \quad (3)$$

Figure 14 plots this metric for IPv4 (top) and IPv6 (bottom). In IPv6, the churn per AS was around 3 updates per origin AS since Jan 2004, except for a period of increased activity seen by some monitors in 2014-2015. After 2015, churn per AS has returned to a lower level of approximately 2 updates per AS for all monitors. In IPv4, except for the AT&T and NTT monitors, this metric was around 6-7 updates per origin AS since January 2003. Other monitors that peer with the Oregon-IX collector show similar behavior. We hypothesize that the anomalies exhibited by the AT&T and NTT monitors are caused by non-stationary periods. In the previous version of this paper [6], we confirmed this intuition by filtering out noise from the AT&T time series as described in [29]; the filtered time series exhibited stable value of updates per origin AS similar to other monitors.

To summarize, while previous work has pointed to the qualitatively different growth trajectories of IPv4 and IPv6 churn, we showed that churn in both protocols grows at the same rate as the underlying topology. We emphasize that understanding the evolution of update dynamics requires examining more than temporal evolution; we must also consider the evolution of the underlying topology. Measuring churn normalized by the size of the underlying topologies reveals a richer picture, namely that *BGP update dynamics in IPv4 and IPv6 are qualitatively similar and their growth is a function of the growth in the number of ASes*. That the average number of updates per AS is 7 in IPv4 and 3 in IPv6 is interesting. In fact, this can be attributed to the fact that IPv4 ASes announce more prefixes on average, and so there are more units that can potentially be updated when an AS becomes active [32].

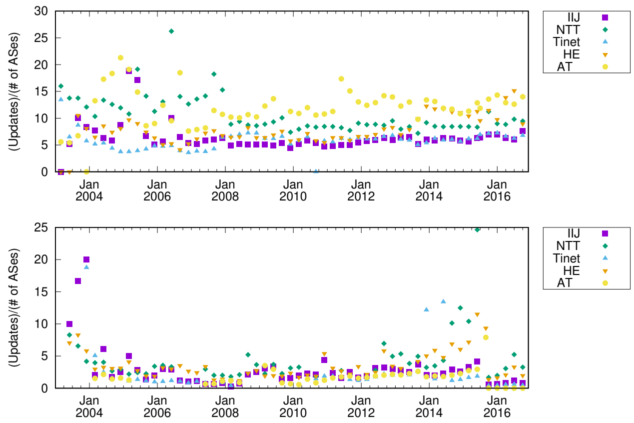


Fig. 14. Churn growth in relation to topology size in IPv4 (top) and IPv6 (bottom). BGP churn, in both IPv4 and IPv6, grows linearly with the number of ASes.

Churn seen from different vantage points: The churn seen from different vantage points can shed some light on the maturity of the underlying topology, because as ASes establish denser interconnections and enforce business relationships, *churn becomes more localized*, i.e., some routing events only affect a limited part of the Internet.

We calculate the cross-correlation between all pairs of daily churn time series in IPv6 and IPv4 respectively, for the monitors in Figure 14. We use the non-parametric Kendall's τ rank correlation coefficient [33], a measure of association between random variables based on the ranking of their sample data. Let $(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)$ be a set of daily churn time series for any two monitors in Figure 14. Any pair of daily churn (x_i, y_i) and (x_j, y_j), where $i \leq j$, are said to be concordant if the ranks for both elements agree: that is, if both $x_i \geq x_j$ and $y_i \geq y_j$; or if both $x_i \leq x_j$ and $y_i \leq y_j$. They are said to be discordant, if $x_i \geq x_j$ and $y_i \leq y_j$; or if $x_i \leq x_j$ and $y_i \geq y_j$. if $x_i x_j$ and $y_i y_j$, the pair is neither concordant nor discordant.

$$\tau = \frac{(n_c) - (n_d)}{n(n-1)/2} \tag{4}$$

where n_c is defined as the number of concordant pairs, and n_d is defined as the number of discordant pairs. Kendall's τ takes values in the range [-1,1]; a value of 1 denotes a perfect correlation, and a value of -1 denotes anti-correlation. Figure 15 shows the calculated correlation coefficients between all pairs of IPv6 time series, as well as between all IPv4 pairs. The x-axis represents a pair of monitors which are, in sequence: IIJ and NTT, IIJ and Tinet, IIJ and HE, IIJ and AT&T, NTT and Tinet, NTT and HE, NTT and AT&T, Tinet and HE, TInet and AT&T, HE and AT&T. We find that IPv6 pairs show strong positive correlation, with τ values ranging from around 0.4 to 0.6. Monitor pairs in general show lower correlation over IPv4 than IPv6. The pairs involving Hurricane Electric (HE) are especially interesting, as their IPv4 τ values are close to 0, e.g., HE and IIJ (-0.023), HE and NTT (-0.024), HE and Tinet (-0.006), and HE and AT&T (0.073). The lower correlation between monitors in IPv4 indicates that churn is highly dependent on the location and configuration of the corresponding router. As stated earlier, this is likely due to denser interconnection and enforcement of business relationships in the IPv4 topology. We have studied this effect in our previous work, where we showed that correlation between IPv4 churn time series doubles after filtering out updates triggered by routing events that affect only a limited part of the Internet [29].

To summarize, we find that the churn seen by different BGP vantage points shows stronger correlation in IPv6 as compared to that in IPv4. Two factors might contribute to the stronger correlation between IPv6 time series than in IPv4. First, the IPv6 AS graph is much smaller and thus provides less isolation (i.e., routing changes will have a larger scope of impact). A second possibility is that since IPv6 deployment is still at an early stage, business policies may be less enforced and monitored, which would also result in less isolation of BGP messages. As IPv6 deployment proceeds, we expect both of these factors to change; the IPv6 topology is growing exponentially, interconnection is becoming denser, and business relationships in IPv6 will start to be enforced. Thus, we expect the correlation of IPv6 churn seen from different BGP monitors to decrease over time and become similar to that in IPv4.

B. Path exploration and Convergence times

Routing changes have different outcomes. Some changes result in the withdrawal (addition) of a prefix from (to) the routing table. Other changes alter the reachability information to a prefix (e.g. rerouting). In addition, routing changes can be transient or long-lasting. The effects of routing instability

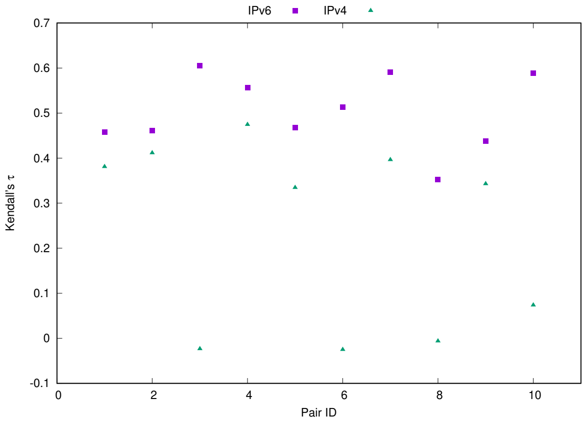


Fig. 15. Correlation of the BGP churn time series across monitors. IPv6 monitors exhibit stronger correlation than IPv4 monitors.

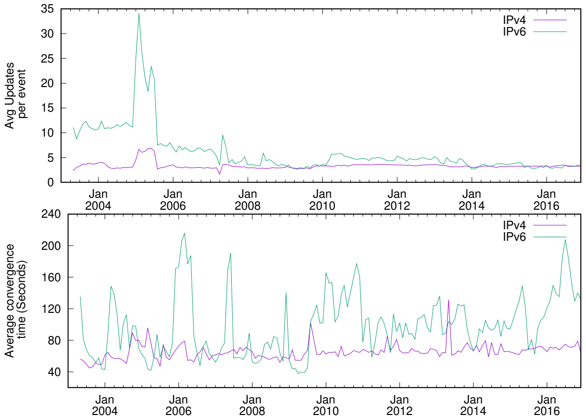


Fig. 16. Comparing path exploration and convergence times (IIJ). The average number of updates per routing change event gradually converges between IPv4 and IPv6. The average convergence time in IPv6 is burstier, with a lower bound at a similar level as in IPv4.

on data plane characteristics such as loss rate have been well studied [26]. It is thus important to compare routing changes in IPv6 and IPv4. But, first we must identify and group prefix updates that constitute a routing change. When an underlying incident triggers a routing change, it often results in several updates for each affected prefix (i.e., convergence sequence). The duration of this convergence sequence is referred to as *convergence time*. A *prefix event* is a sequence of updates for a given prefix that are likely triggered by the same underlying cause. We use the definition by Wu et al. [34] to identify prefix events: *Two consecutive updates for the same prefix belong to the same prefix event if they are no more than 70 seconds apart. The maximum duration for a prefix event is set to 10 minutes. Events with duration longer than 10 minutes are considered to be flapping.* These prefix events can be classified based on the best known path to the affected prefix before and after the event [35], [36]. After identifying all prefix events in our time series, we compute two metrics reflecting their impact: path exploration (average number of updates observed per event) and convergence time. When a route to a prefix fails, BGP may explore several routes before converging to a new route or withdrawing the prefix altogether. A longer path exploration extends BGP convergence time which will likely impede data plane performance.

Path exploration: Path exploration is often more pronounced in events that lead to a complete withdrawal of a prefix (AW events) [36]. The top panel in Figure 16 compares the average number of updates per an AW event in IPv6 and IPv4 as seen from the perspective of IIJ. In IPv4, this number has mostly remained stable below 4. In IPv6, on the other hand, this number was around 10 updates until early 2005, then surged to almost 35 updates and was around 20 updates for 6 months, following which it decreased gradually and since early 2009 has been close to the value for IPv4 at between 3-5 updates. We can imagine two possible causes for this gradual reduction in path exploration. First, ten years ago only a few hundred ASes had deployed IPv6, and routing policies may have been less enforced, allowing exploration of many more alternative paths. Second, the early IPv6 graph was sparser, so paths were naturally longer (see Figure 8), leading to proportionally longer convergence times (i.e., more path exploration) [37]. The trends identified above are consistent across monitors.

Convergence times: The bottom panel in Figure 16 shows the evolution of BGP convergence time from the perspective of IIJ, measured as the monthly average of all prefix event durations (i.e., the time difference between the first update and the last update in an event). The average convergence time in IPv4 is stable around 60 seconds, but is higher and less stable in IPv6. During 2004, IPv6 convergence time was slightly higher, similar to the path exploration metric. We also recorded two periods with sustained higher convergence times, in 2006 and 2010. We found that the increase in 2006 was caused by one prefix that flapped between two paths that only differed in the ATOMIC_AGGREGATE attribute (i.e., one path is announced as aggregated while the other as not). The fact that a single prefix has a large impact on the measured convergence time is surprising. However, our data shows that the small size of the IPv6 routing system makes it vulnerable to such effects. The number of prefix events per day rose sharply from ≈ 200 before this instability to ≈ 350 due to the unstable prefix; this single flapping prefix experienced 150 instabilities per day. On the other hand, the number of prefix events per day is two orders of magnitude larger in IPv4 than in IPv6. When we exclude events related to this prefix, the convergence time drops to the same level as prior to the instability. A similar flapping that involved five prefixes caused the peak in 2010. We believe that this activity was triggered by an IGP misconfiguration in the origin AS. These peaks were evident in all monitors, consistent with the earlier observed strong correlation between IPv6 monitors. In the 5 years since our previous study, the average convergence time in IPv6 has been variable, but consistently higher than that in IPv4.

To summarize, *the evolution of path exploration and convergence time shows that the characteristics of routing changes in IPv6 are gradually becoming similar to those in in IPv4, though convergence times appear to be more variable and larger than those in IPv4. While the similarity between IPv6*

Due to space limitations, we only present results for the IIJ monitor. Other monitors show qualitatively similar results.

and IPv4 in the number of updates per event does suggest a gradual maturity in IPv6 routing, we also find the presence of a few pathologically unstable prefixes in IPv6 that skew the distribution of convergence times.

VII. IPV4 VS. IPV6 PERFORMANCE

Figure 17 plots the relative performance as measured by relative download times for all four vantage points we use (method described at the end of Section 2). In sharp contrast to the measurements of Nikkhah et al. [22] and our own prior measurements [6], we found that IPv6 performance was largely comparable with (or better than) IPv4 performance. In [6], overall, only about 22% of paths in IPv6 performance was better than IPv4 regardless of the paths same or different. In Dec 2016, the IPv6 performance had a qualitative shift from what we saw in 2012. IPv6 performance was comparable with IPv4 right now. When the forward AS-level paths were the same, IPv6 performance was better than IPv4 performance in 49.4% of cases. When the paths were different, IPv6 performance was better in 50.68% of cases. We use a threshold of 10% to determine when performance over the two protocols is “similar”. We found that in 91% of cases IPv6 performance was within 10% of IPv4 (or IPv6 had better performance) if the forward AS-level path was the same in both protocols. In 84% of cases IPv6 performance was within 10% of IPv4 performance (or IPv6 had better performance) if the forward AS-level path was different. However, our measurements are dominated by the path RTT because the transfers are typically small, although we only analyze measurements of transfers over 100K bytes. The dashed lines in Figure 17 plot the relative RTTs measured to the same (IPv4 and IPv6) web-servers; the solid lines (relative fetch time) and the dashed lines (relative RTTs) are closely. When IPv6 performance is better, it is also likely correlated with a same forward AS-level path.

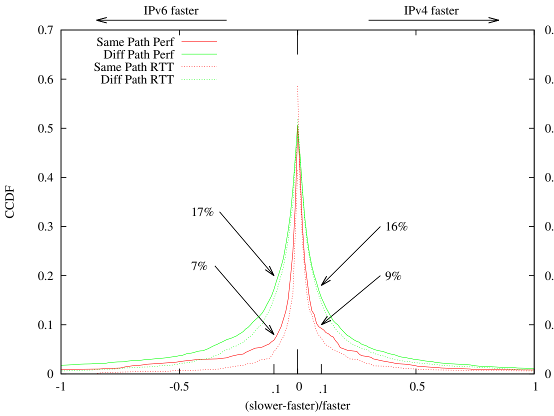


Fig. 17. Relative performance between IPv4 and IPv6 measured by the relative mean fetch times (solid lines) and minimum SYN/ACK RTT (dashed lines). Because small pages (although over 10K bytes) are fetched, performance is dominated by relative RTT. The annotations at x=0.1 represent the points where performance is at least 10% worse in IPv4 or IPv6. IPv4 and IPv6 performance is more likely to be similar if the same AS-level paths are used for both IP protocols.

A. Implications of correlation between performance and topological congruence

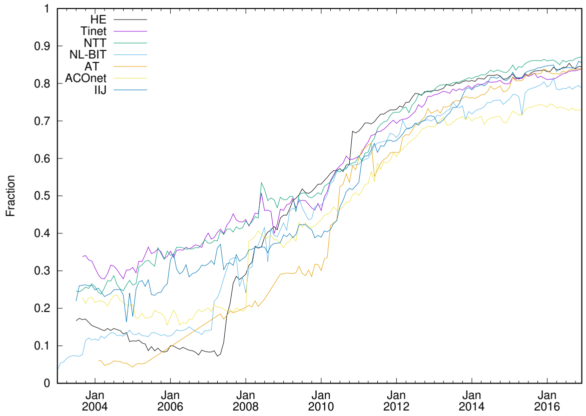


Fig. 18. Fraction of dual-stack ASes reachable using an IPv4 AS path where all AS links in that path are in the IPv6 AS-level graph. If IPv6 BGP paths chosen were consistent with IPv4 paths, then 71%-84% of ASes could be reached over a path congruent in IPv4 and IPv6.

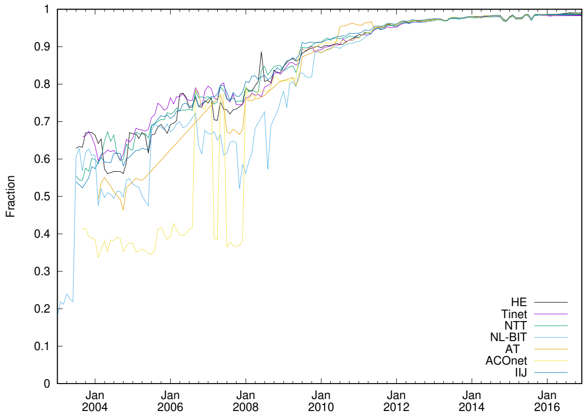


Fig. 19. Fraction of dual-stack ASes reachable using an IPv4 AS path where all ASes in that path are in the IPv6 AS graph. If current IPv6 capable ASes established peerings equivalent in IPv4 and IPv6, then 98% of paths would be identical.

The correlation between data plane performance between two endpoints and the congruity of the AS path between them raises the question: how far are we from topological parity between IPv4 and IPv6? Recall from Section IV that between 40%-75% of AS paths were identical. There are several reasons why the AS paths in IPv4 and IPv6 could be different – the ASes or AS links seen on the IPv4 path may not be present in the IPv6 topology, or networks may simply choose different routes for IPv4 and IPv6. With the data available to us, we cannot confirm why the AS paths differ; we can, however, measure how much congruence between IPv4 and IPv6 AS paths is possible today. For each link in an IPv4 AS path toward a dual-stacked origin AS, we examine whether that link is present in the IPv6 topology, regardless of the AS path on which it appears. Figure 18 shows that by the end of 2016, 71%-84% of AS paths could be identical in IPv4 and IPv6 without creating a new AS links, because for these paths each IPv4 link is already present in the IPv6 topology, just not yet part of an observable AS path between the edges.

We take a step further and examine what would happen if each AS present in IPv6 graph were to establish equivalent peerings in IPv6 and IPv4. Figure 19 shows the fraction of IPv4 AS paths where each AS on the path appears in the IPv6 topology. If ASes present in the IPv6 graph as of December 2016 established equivalent peerings in IPv4 and IPv6, 98% of AS paths could be identical in IPv4 and IPv6, i.e., for an AS link on such a path, both ASes are present in the IPv6 topology, and both ASes already peer in IPv4. If these ASes also started IPv6 peering, we could see the AS paths converge. These results are encouraging, but they are even more motivating when juxtaposed with performance measurements which show that the similarity between IPv4 and IPv6 data performance is correlated with AS path congruity. Together, these results demonstrate the undeniable benefit of BGP peering parity between IPv4 and IPv6 AS-level topologies.

VIII. RELATED WORK

Many attempts have been made to evaluate the status of IPv6 adoption and penetration [38]–[47]. None have found significant activity, even though IPv6 has been implemented on all major network and host operating systems. Google plots a time-series of the percentage of Google users that would access www.google.com over IPv6 if it had an IPv6 address, which moved from 0.6% in May 2012 to 15.6% in Dec 2016 [48]. RIPE NCC shows the percentage of networks (ASes) that announce an IPv6 prefix for all countries from 2.36% in Jan 2004 to 22.72% in Dec 2016 [49]. These measurement studies were either focused on IPv6 *capability*, *i.e.*, how many websites and clients were IPv6 capable, or on the actual levels of IPv6 traffic on the network. In our work, we have focused on IPv6 deployment at the level of organizations, represented in BGP as Autonomous Systems. Huston and Michaelson [50] of APNIC examined a range of types of data collected over four years (January 2004 to April 2008) in search of IPv6 deployment activity. They analyzed inter-domain routing announcements, APNIC's web access logs, and queries of reverse DNS zones that map IPv4 and IPv6 addresses back to domain names. All of their metrics showed some increase in IPv6 deployment activity starting in the second half of 2006, but they emphasized the data's limitations, since it mostly reflected some interest in IPv6 rather than usable IPv6 support. More recently, Czyz *et al.* [51] used a set of 12 metrics to investigate the adoption of IPv6 adoption from 2004 to 2014. They noted that IPv6 adoption as measured by these different metrics varies by up to two orders of magnitude, and shows strong regional differences. In their study, Czyz et al. did not investigate the IPv6 topology, which is the main focus of our work. Michaelson [52] measured the disparity between IPv6 capability at the network level, and IPv6 capability of end-users. Karpilovski *et al.* [53] measured IPv6 deployment using data on address allocation, BGP routing, and traffic. They concluded that even though IPv6 address allocations were increasing, actual traffic levels remained negligible. Huston [54] continuously tracks the evolution of the IPv6 topology and routing, with some high-level comparisons with the current state of IPv4. Aben [55] provides an interactive look into the deployment of IPv6 at the AS-level, further divided by country. To the best of our knowledge, ours is the first work to compare and contrast IPv6 evolution with that of the IPv4 ecosystem.

BGP update dynamics and scalability have been active topics of research during the last decade or so, mostly for the IPv4 topology, *e.g.* [36], [37]. Lately, however, there has been some concern about the *scalability of BGP interdomain routing* [28]. Huston [31] compared update churn in IPv4 and IPv6, and found that while churn in IPv4 does indeed appear "flat", that in IPv6 increases exponentially. In this paper, we compared IPv4 and IPv6 update dynamics, and showed that they are qualitatively similar. The apparent difference between the absolute volume of IPv4 and IPv6 updates over time is simply a function of the different growth rates of the underlying topologies – the IPv4 topology grows linearly, while the IPv6 topology grows exponentially.

A measurement study by Nikkhah *et al.* [22] compared performance (measured in terms of web page download times) over IPv4 and IPv6, with the goal of determining whether the control plane or the data plane was responsible for worse performance over IPv6. They found that while the data plane performs comparably in IPv4 and IPv6, differences in the control plane (routing) are responsible for performance differences seen between IPv4 and IPv6. Huston [56] reported on the comparison between IPv4 and IPv6 performance in 2016, concluding that IPv6 performance is by and large similar to, and often better than, IPv4 performance. Livadariu et al. [57] also compared IPv4 and IPv6 data plane performance, finding that as of 2016, performance between the two protocols was largely comparable. We show in this work that web page download time is dominated by RTT because the pages fetched are typically small, so these performance measurements are dominated by delay rather than available bandwidth. We also demonstrate there is significantly more gain that could be made with the existing ASes that have deployed IPv6; if equivalent links are established in IPv6 as in IPv4 then 98% of existing paths could be identical.

IX. CONCLUSIONS AND FUTURE WORK

With the Internet Assigned Numbers Authority (IANA) and 4 of the 5 Regional Internet Registries (RIRs) now having exhausted their pool of available IPv4 addresses, there has been growing interest in how IPv6 is being deployed and used. Our previous work [6] compared and contrasted the evolution of IPv6 with how IPv4 evolved over the last decade and a half; in this paper we update that study by extending our datasets until the end of 2016. Our findings hint that the IPv6 network is indeed maturing, though the slowdown in IPv6 growth from exponential to linear from 2012 to the end of 2016 is a concerning sign. We found that based on data until the end of 2016, IPv6 adoption was distinctly non-uniform, both topologically and geographically. From the topological perspective, IPv6 deployment was stronger in the core of the network, driven by transit and content providers, while it lagged at the edges, which mostly consist of enterprise customers. While the data at our disposal does not allow us to study *why* deployment is lagging at the edge, we conjecture that this is due to a lack of incentives for edge networks to deploy IPv6, given available alternative strategies, e.g., NAT. A single player, Hurricane Electric, predominated the IPv6 topology significantly more than the most predominant AS in the IPv4 topology. This suggests that several graph-theoretic metrics (e.g., average AS-path length) could be significantly skewed by the single large player in IPv6. In terms of geographical trends, IPv6 adoption was higher in Europe and the Asia Pacific region. We conjecture that adoption in the Asia-Pacific region was spurred by IPv4 address exhaustion, which happened first in that region. A big push toward IPv6 by network operators in the RIPE region could explain why Europe is ahead of North America. From the point of view of routing dynamics, we find that IPv6 behaved mostly like IPv4. Interestingly, we found that while only 40-70% of AS paths were identical in IPv4 and IPv6, up to 71-84% of AS paths could be identical if current IPv6-capable ASes links established equivalent peerings in IPv4 and IPv6, and up to 98% of AS paths could be identical if current IPv6-capable ASes established equivalent peerings.

X. APPENDIX

A. *AFRINIC and LACNIC regional growth*

Figure 20 shows the evolution of the number of IPv4 and IPv6 ASes in LACNIC and AFRINIC, which were omitted from the graph in section III-B.

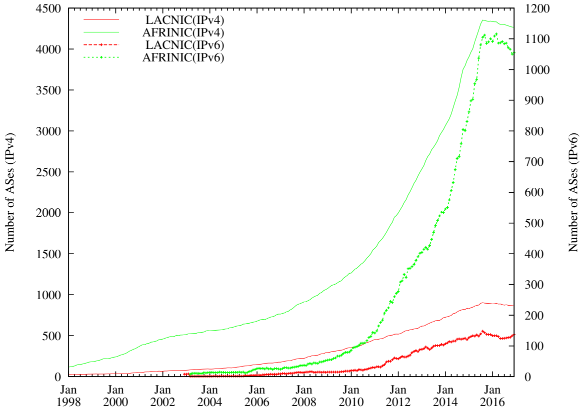


Fig. 20. AFRINIC and LACNIC Regional growth in IPv4 and IPv6 ASes.

B. *Growth by region and business type*

Figure 21 shows the regional growth of different business type classifications in IPv4. Although growth in ECs in different regions mostly follows the same trends as for all ASes (shown in Figure 4), TAs and CPs behave differently. In the IPv4 graph, the growth rate of ARIN-registered TAs was almost identical to that of RIPE-registered TAs (around 7 to 10 ASes/month) until 2002. Since 2002, however, the growth rate of ARIN-registered TAs has slowed to 4 ASes/month, while that of RIPE-registered TAs is around 7 ASes/month. Consequently, the number of RIPE-registered TAs surpassed ARIN-registered TAs in July 2001. In Dec 2016, the number of RIPE-registered TAs was more than twice the number of ARIN-registered TAs. This difference may derive from contrasting regulatory environments which led to more competition in the transit market in Europe than in North America. Another explanation is the tendency of small Eastern European networks to use Provider-Independent (PI) address space [58] which is typically advertised in BGP with its own ASN, rather than Provider-Aggregatable (PA) address space which is typically advertised in BGP by a provider network.We observe that before April 2007 the number of ARIN-registered CPs was larger than the number of RIPE-registered CPs. Since then, the number of RIPE-registered CPs has been larger than ARIN-registered CPs due to their higher growth rate.

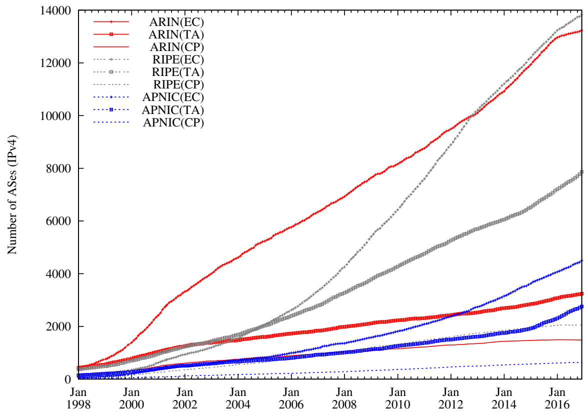


Fig. 21. The regional growth of IPv4 ASes classified by business type. The growth rate of RIPE-registered ECs and TAs is larger than that of ARIN-registered ECs and TAs. Consequently, the number of ECs and TAs is now larger in the RIPE region than in ARIN, though the reverse was true earlier.

Figure 22 shows the growth of IPv6 ASes in different regions classified by business type. As we described earlier (Section III-A), the IPv6 graph is dominated by TAs, and this is true for each of RIPE, ARIN and APNIC. In the IPv6 graph the number of TAs, CPs and ECs in the RIPE region has exceeded that of the ARIN region right since the start of our data collection in 2003, consistent with the stronger community pressure in Europe for operators to support IPv6, including European Commission funding for IPv6 deployment from its early stages.

ACKNOWLEDGEMENTS This work was supported in part by NSF (U.S. National Science Foundation) under CNS-1528148 and CNS-1111449.

REFERENCES

[1] S. Deering and R. Hinden, “RFC 2460. Internet Protocol, Version 6 (IPv6) Specification,” 1998.
[2] “IPv4 Address Report,” http://www.potaroo.net/tools/ipv4/index.html.
[3] Google, “Google ipv6 implementors conference,” 2010, https://sites.google.com/site/ipv6implementors/2010/agenda.
[4] R. Mohan, “Will U.S. Government Directives Spur IPv6 Adoption?” September 2010, http://www.circleid.com/posts/20100929_will_us_government_directives_spur_ipv6_adoption/.
[5] Wikipedia, “IPv6 deployment,” 2012, http://en.wikipedia.org/wiki/IPv6_deployment#Deployment_by_country.

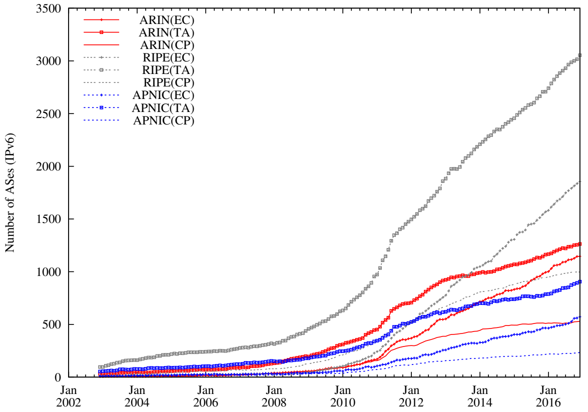


Fig. 22. The regional growth of IPv6 ASes classified by business type. The number of RIPE-registered TAs, CPs, and ECs has always been larger than the corresponding ARIN-registered ASes.

[6] A. Dhamdhere, M. Luckie, B. Huffaker, k. claffy, A. Elmokashfi, and E. Aben, "Measuring the Deployment of IPv6: Topology, Routing and Performance," in *Proceedings of the ACM SIGCOMM Internet Measurement Conference (IMC)*, Nov 2012.

[7] A. Dhamdhere and C. Dovrolis, "Twelve Years in the Evolution of the Internet Ecosystem," *IEEE/ACM Transactions on Networking*, vol. 19, no. 5, 2011.

[8] M. Luckie, B. Huffaker, k. claffy, A. Dhamdhere, and V. Giotsas, "AS Relationships, Customer Cones, and Validation," in *Proceedings of ACM SIGCOMM Internet Measurement Conference (IMC)*, Oct 2013.

[9] David Meyer, "University of Oregon Route Views Project," http://www.routeviews.org/.

[10] RIPE, "Routing Information Service (RIS)," 2008, http://www.ripe.net/ris/.

[11] Q. Chen, H. Chang, R. Govindan, S. Jamin, S. Shenker, and W. Willinger, "The Origin of Power-Laws in Internet Topologies Revisited," in *Proceedings of IEEE Infocom*, 2002.

[12] R. Cohen and D. Raz, "The Internet Dark Matter - On the Missing Links in the AS Connectivity Map," in *Proceedings of IEEE Infocom*, 2006.

[13] Y. He, G. Siganos, M. Faloutsos, and S. V. Krishnamurthy, "A Systematic Framework for Unearthing the Missing Links: Measurements and Impact," in *Proceedings of USENIX/SIGCOMM NSDI*, 2007.

[14] H. Chang and W. Willinger, "Difficulties Measuring the Internet's AS-Level Ecosystem," in *Proc. Annual Conference on Information Sciences and Systems*, 2006.

[15] B. Zhang, R. Liu, D. Massey, and L. Zhang, "Collecting the Internet AS-level Topology," *ACM SIGCOMM CCR*, 2005.

[16] B. Zhang, V. Kambhampati, M. Lad, D. Massey, and L. Zhang, "Identifying BGP routing table transfers," in *Proceedings of ACM SIGCOMM MineNet workshop*, 2005.

[17] "PeeringDB," http://www.peeringdb.com.

[18] "CAIDA - AS Rank," http://as-rank.caida.org.

[19] "Alexa," www.alexa.com.

[20] "The UCSD Network Telescope," https://www.caida.org/projects/network_telescope/.

[21] "CAIDA - AS Classification," https://www.caida.org/data/as-classification/.

[22] M. Nikkhah, R. Guérin, Y. Lee, and R. Woundy, "Assessing IPv6 Through Web Access a Measurement Study and its Findings," in *Proceedings of ACM CoNEXT*, 2011.

[23] M. Luckie, "Scamper: a Scalable and Extensible Packet Prober for Active Measurement of the Internet," in *Proceedings of ACM SIGCOMM Internet Measurement Conference (IMC)*, Nov. 2010.

[24] I. Society, "World ipv6 day," 2011, http://www.worldipv6day.org/.

[25] Hurricane Electric, "Hurricane Electric Peering Policy," http://www.he.net/peering.html.

[26] F. Wang, Z. M. Mao, J. Wang, L. Gao, and R. Bush, "A Measurement Study on the Impact of Routing Events on End-to-end Internet Path Performance," in *Proceedings of ACM SIGCOMM*, 2006.

[27] G. Huston and G. Armitage, "Projecting Future IPv4 Router Requirements from Trends in Dynamic BGP Behaviour," in *Proc. ATNAC*, Dec 2006.

[28] D. Meyer, L. Zhang, and K. Fall, "Report from the IAB Workshop on Routing and Addressing," RFC 4984, 2007.

[29] A. Elmokashfi, A. Kvalbein, and C. Dovrolis, "BGP Churn Evolution: A Perspective from the Core," *IEEE/ACM Transactions on Networking*, vol. 20, no. 99, 2011.

[30] G. Huston, "BGP in 2009 (and a bit of 2010)," Presentation at ARIN XXV meeting, 2010.

[31] Geoff Huston, "The BGP World is Flat," Nov. 2011, http://www.potaroo.net/ispcol/2011-12/flat.html.

[32] A. Elmokashfi and A. Dhamdhere, "Revisiting BGP Churn Growth," *ACM SIGCOMM Computer Communication Review (CCR)*, vol. 44, no. 1, pp. 5–12, Jan 2014.

[33] M. Hollander and D. A. Wolfe, *Nonparametric Statistical Methods*, 2nd ed. Wiley, 1999.

[34] J. Wu, Z. M. Mao, J. Rexford, and J. Wang, "Finding a Needle in a Haystack: Pinpointing Significant BGP Routing Changes in an IP Network," in *Proceedings of USENIX/SIGCOMM NSDI*, 2005.

[35] C. Labovitz, G. R. Malan, and F. Jahanian, "Origins of Internet Routing Instability," in *Proceedings of IEEE INFOCOM*, 1999.

[36] R. Oliveira, B. Zhang, D. Pei, R. Izhak-Ratzin, and L. Zhang, "Quantifying Path Exploration in the Internet," in *Proceedings of ACM SIGCOMM Internet Measurement Conference (IMC)*, 2006.

[37] C. Labovitz, A. Ahuja, A. Bose, and F. Jahanian, "Delayed Internet Routing Convergence," in *Proceedings of ACM SIGCOMM*, 2000.

[38] H. Ringberg, C. Labovitz, D. McPherson and S. Iekel-Johnson, "A One Year Study of Internet IPv6 Traffic," 2008, http://www.nanog.org/meetings/nanog44/presentations/Tuesday/Ringberg_measurement_N44.pdf.

[39] L. Colitti, S. H. Gunderson, E. Kline, and T. Refice, "Evaluating IPv6 Adoption in the Internet," in *Proceedings of PAM*, 2010.

[40] M. Leber, "Global IPv6 Deployment Progress Report," 2006, http://bgp.he.net/ipv6-progress-report.cgi.

[41] M. Kuehne, "Examining Actual State of IPv6 Deployment," 2008, http://www.circleid.com/posts/81166_actual_state_ipv6_deployment/.

[42] M. Prior, "IPv6 survey," 2009, http://www.mrp.net/IPv6_Survey.html.

[43] M. Abrahamsson, "Some Real Life Data," 2008, http://article.gmane.org/gmane.ietf.v6ops/9116.

[44] E. Aben, "IPv4/IPv6 measurements for: RIPE NCC," November 2010, http://albatross.ripe.net/v6-clientresolver.

[45] Emile Aben, "Measuring IPv6 at Web Clients and Caching Resolvers," Mar. 2010, https://labs.ripe.net/Members/emileaben/.

[46] Craig Labovitz, "IPv6 Momentum?" http://asert.arbornetworks.com/2010/10/ipv6-momentum/.

[47] Roch Guerin, "IPv6 Adoption Monitor," http://mnlab-ipv6.seas.upenn.edu:8080/monitor/index.html.

[48] Google, "IPv6 Adoption," https://www.google.com/intl/en/ipv6/statistics.html.

[49] RIPE NCC, "IPv6 Enabled Networks," http://v6asns.ripe.net/v/6?s=_ALL.

[50] G. Huston and G. Michaelson, "Measuring IPv6 Deployment," 2008, http://www.nro.net/news/cisp-ipv6.pdf.

[51] J. Czyz, M. Allman, J. Zhang, S. Iekel-Johnson, E. Osterweil, and M. Bailey, "Measuring IPv6 Adoption," in *Proceedings of ACM SIGCOMM*, Aug. 2014.

[52] G. Michaelson, "Measuring IPv6 at the Network and the Customer Level," May 2012, http://www.potaroo.net/iepg/july2002/.

[53] E. Karpilovsky, A. Gerber, D. Pei, J. Rexford, and A. Shaikh, "Quantifying the Extent of IPv6 Deployment," in *Proc. PAM*, Apr 2009.

[54] G. Huston, "IPv6 Reports," http://bgp.potaroo.net/index-v6.html.

[55] E. Aben, "IPv6 Enabled Networks," http://v6asns.ripe.net/v/6.

[56] G. Huston, "IPv6 Performance Revisited," https://blog.apnic.net/2016/08/22/ipv6-performance-revisited/.

[57] I. Livadariu, A. Elmokashfi, and A. Dhamdhere, "Characterizing IPv6 Control and Data Plane Stability," in *Proceedings of IEEE INFOCOM*, Apr. 2016.

[58] R. Wilhelm, "Members and Number Resources, One Year Later," Feb. 2011, https://labs.ripe.net/Members/wilhelm/members-and-number-resources-1-year-later.