University of Pennsylvania
ScholarlyCommons

Departmental Papers (ESE) | Department of Electrical & Systems Engineering

12-1-2011

Assessing IPv6 Through Web Access - A Measurement Study and Its Findings

Mehdi Nikkhah
University of Pennsylvania, mnikkhah@seas.upenn.edu

Roch A. Guérin
University of Pennsylvania, roch@guerinfamily.org

Yiu Lee
Comcast Corp.

Richard Woundy
Comcast Corp.

Suggested Citation:
Nikkhah, M., Guérin, R.A., Lee, Y. and Woundy, R. (2011). Assessing IPv6 Through Web Access - A Measurement Study and Its Findings. In *Proc. ACM CoNEXT 2011*, December 2011, Tokyo, Japan.

© ACM, 2011. This is the author's version of the work. It is posted here by permission of ACM for your personal use. Not for redistribution. The definitive version will be published in Proceedings of ACM CoNEXT 2011. Email permissions@acm.org

This paper is posted at ScholarlyCommons. http://repository.upenn.edu/ese_papers/602
For more information, please contact repository@pobox.upenn.edu.

Assessing IPv6 Through Web Access - A Measurement Study and Its Findings

Abstract
Transitioning an infrastructure the size of the Internet is no small feat. We are in the midst of such a transition, \ie from IPv4 to IPv6. IPv6 was standardized 15~years ago, but until recently there were few incentives to adopt it. The allocation of the last large block of IPv4 addresses changed that, and migrating to an IPv6 Internet has become more urgent. This migration is, however, still rife with uncertainties and challenges. The goal of this paper is to provide insight into this transition, and possibly make it smoother. The focus is on the ``network,'' and the paper reports on extensive measurements that compare and contrast IPv6 and IPv4. Two important hypotheses, denoted as **H1** and **H2**, were identified and validated. **H1** argues that the IPv6 and IPv4 data planes now perform by and large comparably. In contrast, **H2** points to routing differences as the primary culprit behind occurrences of poorer IPv6 performance. In other words, promoting IPv6 and IPv4 peering parity is probably the single most effective step towards equal IPv6 and IPv4 performance

Keywords
IPv6, adoption, performance, measurements

Disciplines
OS and Networks | Systems and Communications

Comments
Suggested Citation:
Nikkhah, M., Guérin, R.A., Lee, Y. and Woundy, R. (2011). Assessing IPv6 Through Web Access - A Measurement Study and Its Findings. In *Proc. ACM CoNEXT 2011*, December 2011, Tokyo, Japan.

© ACM, 2011. This is the author's version of the work. It is posted here by permission of ACM for your personal use. Not for redistribution. The definitive version will be published in Proceedings of ACM CoNEXT 2011. Email permissions@acm.org

This conference paper is available at ScholarlyCommons: http://repository.upenn.edu/ese_papers/602

Assessing IPv6 Through Web Access A Measurement Study and Its Findings*

Mehdi Nikkhah, Roch Guérin
ESE, University of Pennsylvania
(mnikkhah,guerin)@seas.upenn.edu

Yiu Lee, Richard Woundy
Comcast Corp.
(yiu_lee,richard_woundy)@cable.comcast.com

ABSTRACT

Transitioning an infrastructure the size of the Internet is no small feat. We are in the midst of such a transition, *i.e.,* from IPv4 to IPv6. IPv6 was standardized 15 years ago, but until recently there were few incentives to adopt it. The allocation of the last large block of IPv4 addresses changed that, and migrating to an IPv6 Internet has become more urgent. This migration is, however, still rife with uncertainties and challenges. The goal of this paper is to provide insight into this transition, and possibly make it smoother. The focus is on the "network," and the paper reports on extensive measurements that compare and contrast IPv6 and IPv4. Two important hypotheses, denoted as **H1** and **H2**, were identified and validated. **H1** argues that the IPv6 and IPv4 data planes now perform by and large comparably. In contrast, **H2** points to routing differences as the primary culprit behind occurrences of poorer IPv6 performance. In other words, promoting IPv6 and IPv4 peering parity is probably the single most effective step towards equal IPv6 and IPv4 performance

Categories and Subject Descriptors

C.2.1 [**Computer-Communication Networks**]: Network Architecture and Design—*Distributed networks*; C.4 [**Performance of Systems**]: Design Studies

General Terms

Measurements

Keywords

IPv6, adoption, performance, measurements

*The work of the first two authors was supported in part by a grant from Comcast Corp. and by NSF grant CNS-1116039.

.

1. INTRODUCTION

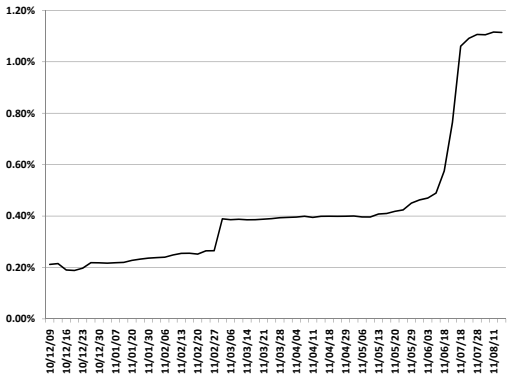


Figure 1: IPv6 Reachability (Top 1M Websites).

The February 3rd, 2011 announcement of the depletion of IANA's free pool of IPv4 addresses[1] brought renewed focus on the Internet's long planned transition to IPv6. This transition; theoretically started 15 years ago when IPv6 was first standardized (RFC 1883 was finalized in 1995), has arguably not been proceeding at a torrid pace. By pretty much any metric, the IPv6 Internet is still dwarfed by its IPv4 counterpart. As of early October 2011, core IPv6 routing tables included less than 7,500 entries, while core IPv4 tables had close to 400,000 (see http://bgp.potaroo.net/index-bgp.html), a more than fifty-fold difference, and barely 10% of all Autonomous Systems (ASes) are announcing IPv6 prefixes (see http://www.ipv6actnow.org/info/statistics/). Similarly, even if as shown in Fig. 1, more than 1% of the top 1 Million[2] web sites are now IPv6 accessible[3], this remains a relatively small fraction of the Internet.

Clearly, this needs to change if IPv6 is to become a viable solution for an impending scarcity of IPv4 addresses. In particular, failure to convince the existing IPv4 Internet to adopt IPv6, *i.e.,* become acces-

[1]See http://www.nro.net/news/ipv4-free-pool-depleted.
[2]According to Alexa's ranking, http://www.alexa.com.
[3]The two "jumps" in Fig. 1 correspond to the depletion of the IANA pool and the World IPv6 Day, respectively.

sible over IPv6, would force the deployment of ever larger translation gateways, *e.g.,* [4], to accommodate the growing numbers of devices provisioned with only an IPv6 address. This would in turn make the transition to IPv6 both complex and onerous. Developing an understanding of current IPv6 adoption, and more importantly of factors that may impede it so as to develop recommendations to remedy them, is what motivated this paper. The transition of an infrastructure of the size of the Internet is, however, a multi-faceted process involving many components, *e.g.,* application software, operating systems, routers, human operators, economic factors, etc., each with its own impact. Assessing the status and role of all of them is neither feasible, nor would it necessarily help pinpoint problem areas. As a result, this study adopts an arguably narrow perspective in assessing the status of IPv6 adoption, and in investigating factors that affect it.

Specifically, we use *access to web content* as the primary vehicle to quantify IPv6 adoption. The Internet obviously comprises many other services, *e.g.,* email, VoIP, streaming, etc., that may or may not be accessible over IPv6 even when web content is[4]. However, access to (web) content is one of the most widespread uses of the Internet, and arguably represents a significant fraction of its "value." Additionally and more pragmatically, developing and deploying large-scale measurements of web access is significantly easier than for other services. In particular, it is easy to obtain a large list of target web sites distributed across the Internet, which can then be regularly monitored for IPv6 accessibility (see Section 3 for details). Hence, offering a reasonably comprehensive view of Internet-wide IPv6 content accessibility.

In addition to focusing on web access, we also limit our investigation to the impact of the *network* on the *quality* of this access. End-hosts, servers, operating systems, and access networks clearly also play an important role. However, besides the difficulty of acquiring local host/server/access information, our main goal is to identify to what extent the IPv6 *Internet* may itself be responsible for a slow IPv6 adoption. In particular, given the lack of meaningful incentives to entice existing IPv4 content providers to become IPv6 accessible[5], avoiding disincentives is vital; and poor connectivity quality would certainly qualify as a disincentive (see [5] for a related discussion). This is readily seen in Google's white-listing process[6], which allows IPv6 connectivity to Google *only* when its quality has been "certified" as on par to that of IPv4. As a result, one of our main concerns is to explore *to what extent* and *why* the quality of IPv6 Internet connectivity may differ from that of IPv4. Hopefully, understanding the why can lead to recommendations on how to remedy deficiencies.

The rest of the paper is structured as follows. Section 2 reviews related measurement studies. Section 3 introduces the system on which the paper's measurement are based. Section 4 introduces the questions the paper seeks to answer, and the methodology it follows. Section 5 presents results obtained from analyzing measurement data, and the paper's main findings and recommendations. Section 6 concludes the paper and points to a number of follow-on investigations.

2. RELATED WORKS

The impending exhaustion of the IPv4 address pool and the many challenges faced by IPv6 have motivated a number of studies. The RIPE Labs IPv6 measurement page at http://labs.ripe.net/content/ipv6-measurement-compilation/ offers a reasonably comprehensive list, and the "Related Links" section of http://mnlab-ipv6.seas.upenn.edu/ points to several other relevant resources. In this section, we briefly discuss a subset of these activities and a few relevant publications.

Geoff Huston's "The ISP Column[7]" offers access to many insightful discussions on topics related to IPv6 (and others). They span both network and end-system deployments, and the problems they have encountered. CAIDA has also carried out various measurement activities dedicated to IPv6 (see kc's blog at blog.caida.org), including measuring IPv6 address allocation and the IPv6 topology.

There have also been several efforts similar in spirit to the one of this paper. Mike Leber from Hurricane Electric used Alexa's top 1 Million web sites to check how many were accessible over IPv6. His focus is limited to IPv6 *accessibility* and his findings are consistent with those of Fig. 1. Mark Prior's IPv6 status survey (http://www.mrp.net/IPv6_Survey.html) has a similar focus, and builds scorecards reporting on IPv6 adoption of several services among academic and research institutions worldwide as well as a number of operators and government organizations. The IPv6 Matrix project (http://www.ipv6matrix.org) developed a tool similar to ours (see Section 3) that crawls through DNS to identify not just IPv6 accessible web servers, but also SMTP and NTP servers. The project shares many of our goals, but has primarily explored geographical differences in IPv6 penetration. Unlike this paper and in spite of the availability of relevant measurement data, it has not really focused on characterizing and explaining performance differences between IPv6 and IPv4.

IPv6 measurements have also been the topic of several papers that we briefly review next.

[4]Mark Prior's ongoing survey of IPv6 accessibility of several "services" (http://www.mrp.net/IPv6_survey.html) gives a partial perspective of those differences.

[5]Many have argued that this is to a large extent why IPv6 adoption has been so slow.

[6]See http://www.google.com/intl/en/ipv6/.

[7]Available at http://www.potaroo.net/ispcol/index.html.

IPv6 penetration was evaluated from multiple perspectives in [6] that looked at address allocation, BGP routes, and traffic volume, and noted that each could give rise to difference conclusions regarding the state of IPv6. More recently, [1] published an extensive one-year study of IPv6 traffic volume across the Internet, though it did not capture tunneled IPv6 traffic.

The issue of IPv6 performance was considered in [2,3, 6, 10, 11]. [2] identified around 8,000 systems reachable over IPv6 and IPv4 from DNS messages, used `ping` to compare IPv6 and IPv4 RTTs, and relied on `traceroute` to investigate the causes of (IPv6) reachability problems. [11] relied on a small set of 26 sites to measure IPv6 and IPv4 one-way delays and delay variations between them, and found that IPv6 performed significantly worse than IPv4 in about 36% of the cases. Poor performing pairs were investigated by comparing their IPv6 and IPv4 paths using again `traceroute`. The paper found that tunnels were often responsible for poor IPv6 performance, and that IPv6 and IPv4 commonly used different paths, with IPv6 routing exhibiting more problems. [10] performed similar measurements but for 936 websites located in different countries, and reached somewhat different conclusions, *e.g.,* it did not identify tunnels as a major source of performance degradations. Finally, [3] evaluated IPv6 performance (and adoption), but from the point-of-view of a *single* website operator, *i.e.,* a perspective symmetric to ours that considers access to many (all) IPv6 accessible websites.

In summary, in spite of a common "theme," namely, the desire to assess IPv6 penetration and performance, this paper differs from these previous works on multiple accounts. The first is the scale of the measurements carried out, *i.e.,* from multiple vantage points, to several millions of web sites, and over a period of many months[8]. The second and possibly more important difference is the focus on using the measurement data together with AS-path information to understand and characterize when and why performance differences exist between IPv6 and IPv4. Identifying what may be causing IPv6 to under-perform together with possible remedies are important contributions of the work.

3. TOOLS AND MEASUREMENTS

Characterizing IPv6 adoption through access to web content calls for **(i)** identifying a sufficiently large and representative set of web sites to monitor; **(ii)** checking IPv6 accessibility of these sites, and comparing web access performance over IPv4 and IPv6 for sites accessible over both; and **(iii)** repeating step (ii) from multiple vantage points across the Internet to avoid biases associated with individual locations[9].

Step **(i)** is based on the top 1 Million (top 1M) web sites list maintained by Alexa (http://www.alexa.com). Alexa's list is retrieved before each round of monitoring, and new sites that have not been seen before (their first appearance in the top 1M) are added to the monitoring list and tracked from this point onward[10]. Step **(ii)** relies on a dedicated monitoring tool, whose structure is shown in Fig. 2. The tool is written primarily in java, and is multi-threaded so that multiple sites (no more than 25 to avoid bandwidth and processing bottlenecks) can be monitored in parallel.

Following Fig. 2, each round of monitoring starts by retrieving the latest top 1M list from Alexa as well as any additional sites that the user wants to manually import. Sites not present in previous monitoring rounds are added to the list of sites that will be monitored from this point onward. Results of each monitoring round are stored in several tables in a `mysql` database. The monitoring of a site is assigned to an individual thread, and proceeds in several phases with the outcome of each phase determining the subsequent one.

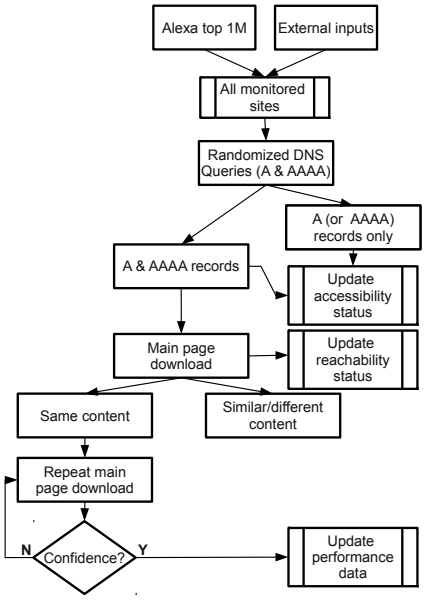


Figure 2: Monitoring software structure.

The first monitoring step involves a DNS query for the A and AAAA records of the site. The order in which sites are monitored is randomized in each round to avoid time-of-day biases in the results. The results of the DNS queries are recorded and sites that have both A and AAAA records proceed to the phase that involves downloading a copy of the site's main page over both

[8]In that respect, the IPv6 Matrix project is the one that comes closest to ours.

[9]Throughout the paper, *site* denotes a web site we monitor for IPv6 and IPv4 accessibility, while a *vantage point* refers to a location from which monitoring is performed.

[10]After less than a year of approximately bi-weekly monitoring rounds, churn alone in Alexa's top 1M list resulted in over 2 millions sites being monitored.

IPv4 and IPv6. Only the main page is retrieved and embedded objects are not. The IPv4 and IPv6 pages are then compared to determine if they are "identical." For the sake of simplicity and expediency, this is done by comparing page sizes, with pages declared identical as long as their byte counts are within 6% of each other (the vast majority of sites have identical IPv4 and IPv6 pages, and the answer is relatively insensitive to the exact threshold used to claim "identity"). The outcome of this first query is recorded and sites with identical IPv4 and IPv6 pages proceed to the next phase. It involves multiple page downloads, first for IPv4 and then IPv6, each after proper resetting to avoid local caching effects. Downloads repeat until the measured average download time is within 10% of the mean with 95% confidence, at which point the page size and its average download time are recorded. After results have been obtained for both IPv4 and IPv6, the thread is returned to the thread pool, and is available to start the monitoring of the next site, if any.

The multiple downloads and associated confidence intervals are to minimize transient fluctuations during a monitoring round. Variations across rounds are, however, still possible; especially since monitoring times are randomized. As a result, a site's "overall" performance is obtained by averaging samples gathered over many months, and by again ensuring that the 95% confidence interval is within 10% of the mean. Sites that do not meet this criterion are not included in the analysis. Possible causes for sites failing this confidence target are discussed in Section 5.1, which also establishes that the removal of those sites does not bias the analysis.

The measurements of Fig. 2 are repeated from multiple vantage points (see Table 1), with each vantage point maintaining a local database of its own monitoring results. A common repository at Penn aggregates the measurement data from the different vantage point.

To fulfill our goal of correlating performance measurements with network characteristics, we also gather AS-level paths from each monitoring vantage-point to all the websites it monitors. This is done by accessing the (core) routing table of a router close to the machine running the monitoring software. For example, for the monitor located at Penn, we had access to the BGP routing tables of one of the routers in the GigaPoP connecting Penn to the rest of the Internet.

The choice of AS-level paths (from BGP routing tables), as opposed to IP-level paths (from `traceroute`), to characterize network connectivity was motivated by several considerations. The first is that comparing IP-level IPv4 and IPv6 paths is challenging, if not impossible. It requires identifying if and when the IPv6 and IPv4 (interface) addresses returned by `traceroute` (to a common destination) map to the same devices. Unfortunately, many of these addresses, *e.g.,* routers' interface addresses, are not registered with DNS, which makes performing such a mapping problematic. This is further complicated by the possible use of tunnels[11] in IPv6 paths (to cross IPv4-only islands). In addition, our initial experiments using `traceroute` to obtain path information were unsuccessful (did not complete) over 50% of the time. This has been reported by others [7], and in our case may have been exacerbated by the large number of destinations being targeted. Finally, even if IPv6 (and IPv4) AS-level and IP-level paths need not always agree, both our own initial tests using `traceroute` and results from earlier studies [7] indicate that while discrepancies exist they are relatively rare. The impact of these few discrepancies is further mitigated by the fact that we are not asking for perfect agreement between AS-level and IP-level paths. Instead, we only look for AS-level agreement between IPv6 and IPv4 as a predictor for IP-level agreement.

4. METHODOLOGY

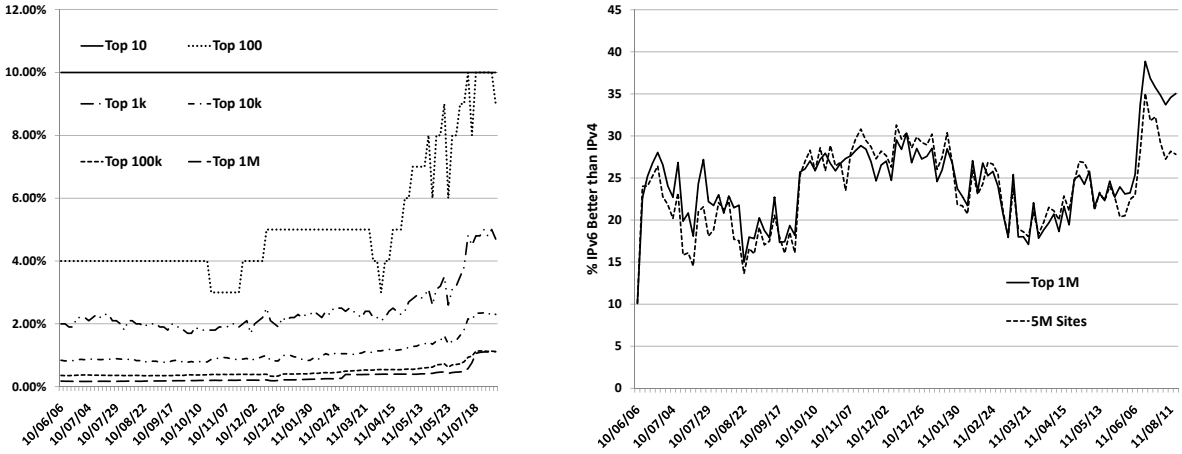
As discussed in Section 3, the list of monitored websites is based primarily on Alexa's top 1M websites. In spite of its size, the use of this list might introduce a bias in the selection of monitored sites. As shown in Fig. 3a, a site rank does influence its likelihood of IPv6 accessibility. However, Fig. 3b illustrates that when comparing IPv6 and IPv4 web access performance[12], there is little difference between Alexa's top 1M websites and another sample of about 5 millions sites obtained by supplementing Alexa's list with (new) sites extracted weekly from Penn's main DNS cache. The figure reports on one aspect of performance, namely, the odds that IPv6 outperforms IPv4, but similar findings held for other metrics. This offers reasonable evidence that conclusions drawn from measurements for the top 1M websites extend to a broader representation.

Avoiding bias when it comes to vantage points selection is more difficult. Nevertheless, as shown in Table 1, the vantage points from which we were ultimately able to obtain data offer reasonable geographic distribution, and include both academic and commercial entities. The first two columns of the table identify the vantage point location, and when its monitoring started. The third column specifies whether or not AS_PATH information was available from that vantage point, while the fourth column reveals if it was white-listed (W-L) by Google. Finally, the last column gives the vantage point "type," *i.e.,* academic or commercial. Vantage points performed one or more rounds of monitoring each week, but were not synchronized.

Data was collected from each vantage point for IPv6

[11] See http://ipv6blog.net/ipv6-tunnel-brokers/ for a list of current IPv6 tunnel brokers.

[12] Unless otherwise stated, in the paper we use download speed as the measure of web access performance.



(a) IPv6 reachability by rank

(b) How often is IPv6 download faster

Figure 3: Influence of site rank on IPv6 reachability & performance (12 months window from Penn).

Vantage Point	Date	AS_PATH	W-L	Type
Comcast (Denver, CO)	2/4/11	Y	N	Comml.
Go6-Slovenia (Slovenia)	5/19/11	N	N	Comml.
Loughborough U. (Great Britain)	4/29/11	Y	N	Acad.
Penn (Philadelphia, PA)	7/22/09	Y	N	Acad.
Tsinghua U. (China)	3/22/11	N	N	Acad.
UPC Broadband (Netherlands)	2/28/11	Y	Y	Comml.

Table 1: Monitoring vantage-points.

and IPv4 web access performance for sites in its list of monitored sites (see Fig. 2). Because monitoring rounds are not synchronized across vantage points, they sample Alexa's top 1M list at different times. This together with different start dates (and in some cases, *e.g.,* Penn, the external input of additional sites), resulted in each vantage point monitoring a slightly different set of sites.

Because of our goal to connect differences between IPv6 and IPv4 web access to network characteristics, we focused on vantage points for which *both* AS_PATH and performance information were available, *i.e.,* vantage points with a "Y" in the AS_PATH column of Table 1.

Several factors can influence web access performance:

(E). The client **E**nd-system;
(S). The **S**erver end-system, including its location;
(D). The network **D**ata plane (packet forwarding, congestion, tunnels, etc.);
(C). The network **C**ontrol plane, as embodied in peering and routing choices.

The client end-system performing the monitoring at each vantage point is a device we control. In most instances[13], monitoring clients ran a version of the Linux operating system on a multi-core CPU with plenty of memory, and had high quality native IPv6 (and IPv4) connectivity. Individual tests on those machines did not identify noticeable performance differences between IPv6 and IPv4, which allows us to eliminate (**E**) as causing differences between IPv6 and IPv4 web access performance from each vantage point. Because our focus is on understanding to what extent the *network* is responsible for differences in performance, we would ideally like to eliminate (**S**) whenever possible. Unfortunately and as mentioned earlier, without direct access to the servers themselves, which we did not have, explicitly isolating the performance impact of those systems is difficult. Eliminating (**S**) as a possible cause for performance differences between IPv4 and IPv6 relies, therefore, on a mostly indirect approach on which we expand later in this section.

The general methodology used to identify possible causes for performance differences between IPv6 and IPv4 web access is broadly outlined in Fig. 4. For each vantage point, web sites are first partitioned in two disjoint sets: Same location (SL) and different locations (DL), where location refers to the AS in which the web site is located. Sites in SL have IPv6 and IPv4 addresses mapped to the same AS, while they were in distinct ASes for sites in the DL set.

Sites in DL are often, though not always, CDN users as most CDN providers do not yet offer production-level IPv6 services, *e.g.,* see http://www.akamai.com/ipv6. Sites in DL are analyzed separately, as their different locations typically imply altogether different network paths that can, therefore, not be readily compared.

[13]The one exception was the Loughborough U. monitor that used a mac-server running Darwin.

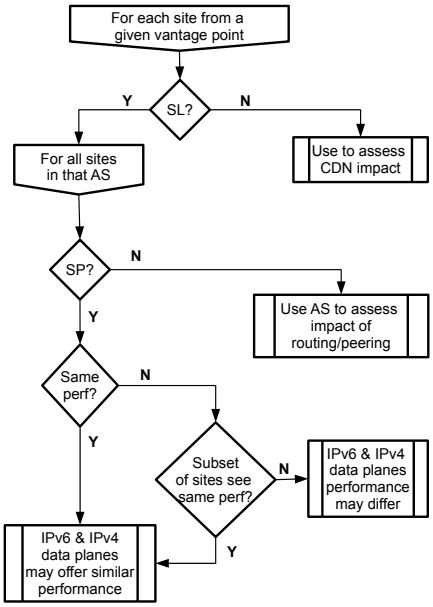


Figure 4: Destination AS classification.

Sites in SL are of greater interest, as they share start and end-points, which makes comparing IPv6 and IPv4 performance more meaningful. The approach we take to identify factors that may be behind differences between IPv6 and IPv4 is structured to test the validity of two hypotheses, which emerged from initial processing of the data. The first hypothesis, denoted **H1**, is as follows:
H1: *The IPv6 data plane performance is mostly on par with its IPv4 counterpart.*

We note that by data plane performance, we refer to IP packet forwarding, and do not imply that tunnels are not present or have no impact[14]. However, as we discuss next, tunnels are unlikely to be present in the primary data set that we use to validate **H1**.

Specifically, to validate **H1** we focus on sites whose IPv6 and IPv4 AS paths coincide so that a side-by-side comparison is meaningful. For each vantage point, the destination ASes of sites in the SL set of that vantage point are, therefore, split in two sub-groups: ASes reachable over the *same* IPv6 and IPv4 paths (SP), and ASes reachable over *different* paths (DP).

For destination ASes in the SP group, because IPv6 and IPv4 make consistent routing decisions, (**C**) can be eliminated as a possible cause of performance differences and only (**S**) and (**D**) remain. IPv6 and IPv4 performance of ASes in SP are, therefore, checked to see if they are comparable (or if IPv6 is better). Our definition of comparable performance is that the average download speeds (across sites in the AS) do not differ by more than 10%; the range of our confidence interval. If performance is comparable, then one can reasonably assume that neither (**S**) nor (**D**) contribute substantially to differences between IPv6 and IPv4 (poor IPv6 data plane performance could arguably be offset by better server performance, but besides low odds for this to be consistently true across sites/servers in an AS, most server benchmarking data [8,9] point to IPv6 server performance at best similar to that of IPv4).

When performance is not comparable, it is still possible for (**D**) not to be responsible for the AS-level differences that are observed, *i.e.,* they can be caused by poor IPv6 support in a majority of servers for sites in that AS. For that purpose, the distribution of IPv6 and IPv4 performance (average download speeds) differences across individual sites in the AS is evaluated.

The presence of a "mode" around zero for that distribution would indicate that some sites experience similar IPv6 and IPv4 performance. Hence, the fact that IPv6 and IPv4 share the same path, would support the claim that (**D**) is not responsible for the difference in performance seen at the AS level (across all sites). Instead, those differences are likely caused by (**S**), *i.e.,* because of a large subset of sites with servers exhibiting poor IPv6 performance. In order to detect the presence of a zero-mode, differences in performance between IPv6 and IPv4 are computed for all sites in an AS. A zero-mode is claimed, if there is at least one site for which this difference is within 10% of IPv4 performance (10% is our measurement confidence target). In most instances of a zero-mode, we had more than 10 sites with comparable IPv6 and IPv4 performance, and it was only for ASes with few sites that zero-modes consisted of only one or two sites. We report in Section 5 on the results of this analysis, as well as additional tests not shown in Fig. 4 that seek further validation through cross-checks across vantage-points when feasible.

ASes (and associated sites) in the DP group are used next to assess the impact of different IPv6 and IPv4 routing choices, *i.e.,* (**C**). Unlike ASes in SP, those in DP are reached over *different* IPv6 and IPv4 paths. Hence making (**C**) a prime candidate to explain performance differences. Clearly, (**S**) and (**D**) are also possible, but assuming that **H1** can indeed be shown to hold, (**D**) can be eliminated so that only (**C**) and (**S**) remain. This makes ASes in DP good candidates to evaluate the impact of routing choices, *i.e.,* (**C**). This is reflected in hypothesis **H2**, which we state next.
H2: *Differences in routing choices between IPv6 and IPv4 are a major cause of poorer IPv6 performance.*

In other words, (**C**) is the major culprit for poor(er) IPv6 performance. The first step in establishing **H2**'s validity is to directly compare the performance of ASes in SP and in DP. The primary difference between the SP and DP groups being (**C**), such a comparison is an

[14]In addition to [10] mentioned earlier, others, *e.g.,* http://mobitech.ie.cuhk.edu.hk/projects/ipv6/tunnel.html, have reported that the performance impact of IP-in-IP tunnels is small.

obvious candidate. Section 5 offers evidence that the results of this comparison are consistent with **H2**. Next, we describe additional steps aimed at refining, whenever possible, the analysis and the extent to which **H2** holds.

We first attempt to rule out any impact from (**D**) on the performance of DP ASes by exploiting SP data gathered across vantage point. Specifically, consider the ASes present on an IPv6 AS path to a destination AS in DP. Our goal is to establish that the data plane, *i.e.,* (**D**), of these ASes (or as many of them as possible) does not contribute to worse IPv6 performance. For that purpose, we check whether they can be found on a "good" IPv6 AS path, *i.e.,* to a destination AS in SP for which IPv6 and IPv4 performance was found to be comparable. If the answer is positive, then that AS can be argued *not* to contribute significantly (through (**D**)) to performance differences between IPv6 and IPv4. If it did, all paths traversing it would be affected. The more "good" paths an AS is in, the stronger the argument. Conversely, the more such ASes we can identify, the less likely it is that (**D**) is contributing to performance differences we may be observing for ASes in DP. In particular, if we know that *all* ASes in a DP AS path are good, then (**D**) is unlikely to be the cause of poorer IPv6 performance we may observe on this path.

The next step is similar in spirit, and seeks to rule out server impact, *i.e.,* (**S**), whenever possible. It focuses on sites located in destination ASes found to be in SP from some vantage-points and in DP in others. As discussed earlier, sites associated with the zero-mode of the distribution of performance differences between IPv6 and IPv4 in such ASes have both network *and* servers that perform equally well in IPv6 and IPv4. Hence, for any AS that is in SP from some vantage point, we can extract the set of sites with servers known to perform well in IPv6. If that AS is in DP from some other vantage point, comparing its IPv6 and IPv4 performance using only those sites allows us to eliminate (**S**) as the cause for any performance difference. Hence, strengthening the case that (**C**) is the main reason.

Note, however, that the step we just described need not be feasible for all ASes, *i.e.,* some ASes may never be in SP from any vantage point, or even fully accurate, *i.e.,* recall that not all vantage points monitor the exact same set of sites, so that sites with servers that perform poorly in IPv6 may be present from one vantage point and not from another. As a result, we perform an additional test for ASes in DP, namely, as with ASes in SP, we look for the possible presence of a zero-mode in the distribution of performance differences between IPv6 and IPv4. The existence of such a mode would indicate that for at least some sites, the IPv6 and IPv4 networks perform comparably, even if they rely on distinct paths. As we shall see in Section 5, there are actually very few ASes for which this holds true.

5. ANALYSIS AND RESULTS

This section reports on results obtained by processing monitoring data obtained from the vantage-points listed in Table 1. Table 2 provides general statistics on the coverage, in terms of number of websites and ASes, for the monitoring from each of the vantage-points[15] for which AS_PATH information was available.

Numbers of	Penn	Comcast	LU	UPCB	All
Sites (total)	12385	4568	5069	7843	NA
Sites kept	7994	3525	3906	4418	NA
Dest. ASes (IPv4)	1047	724	801	766	1364
Dest. ASes (IPv6)	727	592	642	609	1010
ASes crossed (IPv4)	1332	922	1019	988	1785
ASes crossed (IPv6)	849	742	764	746	1208

Table 2: Monitoring profiles per vantage-point.

The first row of the table gives the number of sites accessible over both IPv6 and IPv4 from each vantage-point. As mentioned earlier, differences in monitoring start times, together with non-synchronized sampling of Alexa's top 1M list, and the occasional inclusion of additional sites, *e.g.,* from Penn DNS cache, contributed to vantage-points monitoring different sets of sites, although with a large intersection (of around 1 million).

As mentioned in Section 3, only websites whose statistics met a target confidence were kept. The second row of the table reflects the result of this pruning.

The third and fourth rows give the number of destination ASes from each vantage point and across all vantage points. Sites in the DL category and the use of mechanisms such as '6to4' (RFC 3056), contribute to the differences between IPv6 and IPv4.

Finally, the fifth and sixth rows provide the total number of ASes crossed (including destination ASes) by IPv6 and IPv4 paths; again from each vantage point and across all vantage points. As expected given the relatively low adoption of IPv6 (see Fig. 1), these ASes do not span the entire Internet that has over $37,000$ ASes. Nevertheless, the numbers point to a "reasonable" coverage. We also note that the number of ASes crossed in IPv6 is lower than in IPv4. This is consistent with findings by others (*e.g.,* CAIDA's IPv6 topology monitoring effort) that report a smaller and sparser IPv6 topology, and may also be attributable to tunnels.

5.1 Data Sanitization and Classification

As shown in the first two rows of Table 2, data from a non-trivial number of sites ended-up not being used,

[15]LU stands for Loughborough University.

because they failed to meet target confidence levels. Table 3 explores the reasons behind those failures.

	Insufficient Samples	↑	↓	↗	↘
Penn	2807	180	103	732	569
Comcast	251	83	52	530	127
LU	258	49	63	419	374
UPCB	1146	233	214	1033	799

Table 3: Causes of confidence target failures.

The table identifies five types of causes for a site's performance measurement not achieving the desired level of confidence. In many cases (first data column), the reason is simply that we have not accumulated enough samples to reach the desired accuracy. This was to some extent exacerbated by the recent (June 2011) large jump that IPv6 penetration experienced because of the World IPv6 Day (see Fig. 1). Many new sites were added, and given our weekly rate of monitoring, the number of samples gathered for those sites is still small. The next two columns of Table 3 (vertical up and down arrows) identify two other causes, namely, sites that experienced a sharp upward or downward transition[16] in performance during the measurement period. In some of those cases (64 out of 283 for Penn, 64 out of 135 for Comcast, 43 out of 112 for LU, and 169 out of 447 for UPCB), this transition was the result of a path change, but not in all cases. For those last cases, it may have been the result of equipment upgrades or other causes we cannot infer. The last two columns of the table give the number of sites for which a linear regression revealed a steady upward (downward) trend in performance. We don't have explanations for those changes, but they indicate that performance at those sites is non-stationary and, therefore, cannot be really used to quantify average performance over time.

Next, as discussed in Section 4, data from sites that were kept, was used to classify for each vantage point sites and destinations ASes into DL, SL/SP, and SL/DP categories. Table 4 identifies the number of sites in each category from the vantage points for which both monitoring data and AS_PATH information was available.

	Penn	Comcast	LU	UPCB
# DL sites	784	450	352	485
# SP sites	424	1113	2291	2597
# DP sites	6786	1962	1263	1336

Table 4: Sites classification.

[16]Transitions were detected using a median filter of length 11 configured to report changes in performance of magnitude greater than 30%, *i.e.,* it triggered after 6 or more consecutive samples 30% higher (lower) than the previous ones.

	Penn	Comcast	L.U.	UPCB
SP good perf.	64	185	462	1242
SP bad perf.	8	64	42	163
DP good perf.	404	346	206	463
DP bad perf.	880	93	106	216
DL good perf.	111	54	65	103
DL bad perf.	117	50	24	92

Table 5: Classification of removed sites.

As discussed in Section 4, data from sites in the different categories of Table 4 are used to validate hypotheses **H1** and **H2**. A natural question is, therefore, whether the "removal" of sites that did not meet confidence targets introduced a bias in the data. Table 5 reports on the result of this investigation[17] for sites for which sufficient samples were available, *i.e.,* the last four columns of Table 3. The first two data rows indicate that more SP sites with good performance were removed than with bad performance. If anything, this introduces a bias against the validation of **H1**, which calls for equal IPv6 and IPv4 performance for SP sites. The situation is more nuanced for DP sites. More sites with bad (IPv6) performance are removed at Penn, but the opposite holds at other vantage points. Since DP sites are used to establish that differences in routing are responsible for worse IPv6 performance (hypothesis **H2**), the removal of more "good" DP sites may slightly bias the results in favor of **H2**. However, we note that the difference in the number of good and bad sites being removed is small compared to the total number of sites used to evaluate **H2** (see Table 2). Hence, any bias is weak at most. Finally for completeness[18], the last two rows of Table 5 indicate that more or less the same number of good and bad DL sites are removed from each vantage point.

	Penn	Comcast	LU	UPCB
# sites	784	450	352	485
IPv4≥ IPv6	96%	91%	94%	90%
IPv4 perf.	35.6	49.3	50.9	49.6
IPv6 perf.	28.2	43.6	43.4	47.3

Table 6: IPv6 vs. IPv4 performance (kbytes/sec) for sites in DL.

5.2 Impact of Location and Path Length

This section focuses on two aspects. The first is that of different IPv4 and IPv6 locations, *i.e.,* for sites in DL, and the second is that of IPv6 and IPv4 performance as a function of (AS) path length. The first aspect can provide some insight into the potential benefits (for IPv6)

[17]Performance refers to that of IPv6 relative to IPv4.

[18]They are not used to validate either **H1** or **H2**.

		1 Hop	# sites	2 Hops	# sites	3 Hops	# sites	4 Hops	# sites	> 5 Hops	# sites
Penn	IPv4	25.4	5	39.5	4327	31.1	2318	28.5	567	22.7	179
	IPv6	-	0	104.0	6	33.9	742	28.7	3296	22.1	3352
Comcast	IPv4	57.3	85	42.8	825	39.3	1348	29.8	103	22.8	8
	IPv6	37.2	49	47.1	730	36.0	1302	26.1	159	44.1	129
LU	IPv4	113.3	153	69.8	887	49.0	478	42.8	93	21.4	24
	IPv6	43.4	130	67.2	983	45.3	375	51.5	142	27.0	5
UPCB	IPv4	64.8	2	67.1	194	49.9	1069	47.5	397	34.4	89
	IPv6	-	0	34.5	132	48.7	1071	50.6	485	36.2	63

Table 7: DL+DP sites – Performance (kbytes/sec) by hop count.

of commercial-grade IPv6 CDN offerings. The second aspect is a precursor to the analysis of the next section, which seeks to establish that the IPv6 and IPv4 data planes perform similarly. Table 6 sheds some light on the first issue by comparing IPv6 and IPv4 performance for DL sites from each vantage point. As seen from the table, IPv4 is as good or better as IPv6 a majority of the time, and its average download speed is consistently higher (for DL sites) from all vantage points. This provides some indication of the magnitude of the improvements that native IPv6 CDNs would result in.

Table 7 reports on IPv6 and IPv4 performance for sites in DL and DP, *i.e.,* sites whose IPv6 and IP4 paths differ, as a function of path length (in AS hops). The table provides both the number of sites (from each vantage point) with an AS path of a given length, and the corresponding average IPv6 and IPv4 performance. For small hop-count values (one or two hops), IPv6 and IPv4 performance exhibit differences, with IPv4 usually performing better. A likely explanation[19] is the presence of tunnels that make IPv6 paths appear shorter than they really are. Given that as seen from the table, *i.e.,* from the IPv4 performance rows, performance typically decreases with hop-count, the lower performance of IPv6 paths could be explained by their true hop-count being actually higher. This explanation appears to be borne by the fact that as hop-count increases and tunnels become arguably less likely, the performance of IPv6 and IPv4 becomes comparable. This would in turn support hypothesis **H1**, and the next section offers additional evidence to that effect.

5.3 Identical IPv6 and IPv4 Paths

This section focuses on sites that were reachable over the same IPv6 and IPv4 AS paths, and in particular seeks to confirm hypothesis **H1** that the IPv6 data plane is on par with its IPv4 counterpart. Note that paths in the SP set will most likely not involve tunnels, except possibly internal to an AS.

Table 8 offers data in support of the fact that when IPv6 and IPv4 follow the same path, performance will typically be similar (IPv6 performance is within our 10% confidence interval of IPv4 performance, or is better). The columns of the table correspond to different vantage-points, while its rows report the results of the different processing steps discussed in Section 4.

	Penn	Comcast	LU	UPCB
IPv6≈IPv4	81.3%	80.7%	70.2%	79.8%
Zero mode	9.4%	6%	10.8%	7.3%
Small Number of sites	9.3%	13.3%	19.0%	12.9%
# ASes	75	233	248	124
x-check (+)	47	129	164	82
x-check (−)	0	0	0	0

Table 8: IPv6 vs. IPv4 for SP destination ASes.

Specifically, the first row identifies the percentage of destination ASes for which IPv6 and IPv4 perform similarly (or IPv6 was better) across sites in that AS. The second row identifies the percentage of ASes with lower overall (across sites) IPv6 performance, but for which the distribution of performance differences between IPv6 and IPv4 across sites exhibits a zero-mode. As discussed in Section 4, a zero-mode corresponds to individual sites in the AS experiencing the same IPv6 and IPv4 performance, which argues for the network not being responsible for the lower AS-level IPv6 performance. The next row indicates that the remaining ASes for which IPv6 performed worse than IPv4 and no zero mode could be identified, all included only a "small" number of sites (less than four). Arguably, this small number may be why we were not able to find sites with servers that perform well in IPv6, and which would therefore have contributed to a zero mode in the AS. In any case, the relative number of such ASes is small, so that the data reported in Table 8 broadly supports the conclusion that the IPv6 and IPv4 data planes perform comparably.

Further validation of this conclusion and, therefore, of hypothesis **H1**, is provided in the last two rows of Table 8, which report on the results of cross-checks performed across vantage points. Specifically, results for ASes in SP from different vantage points were com-

[19] We could verify it in a handful of cases but not all.

		1 Hop	# sites	2 Hops	# sites	3 Hops	# sites	4 Hops	# sites	> 5 Hops	# sites
Penn	IPv4	-	0	-	0	36.0	23	29.5	203	29.1	169
	IPv6	-	0	-	0	34.4	23	27.6	203	29.5	169
Comcast	IPv4	64.2	137	41.6	632	36.0	304	36.8	10	-	0
	IPv6	59.9	137	42.1	632	35.4	304	34.0	10	-	0
LU	IPv4	60.3	229	62.5	1829	42.7	115	21.3	16	-	0
	IPv6	57.3	229	62.2	1829	39.2	115	19.4	16	-	0
UPCB	IPv4	-	0	43.7	168	62.8	2202	50.3	38	13.4	1
	IPv6	-	0	41.4	168	64.7	2202	47.6	38	13.7	1

Table 9: Destination ASes in SP: Performance (in kbytes/sec) by hop-count.

pared. A positive cross-check corresponds to an AS being found in the same category (row) of Table 8 from all its vantage points, while a negative cross-check arises if it is found to belong to different categories. Because not all destination ASes can be found to be in SP from more than one vantage point, cross-checks cannot be performed for all ASes. However, for those ASes for which cross-checks were available, the last two rows of Table 8 indicate that all were positive and none were negative. This further strengthens the conclusion that based on the available data, **H1** holds.

Additional support for **H1** is presented in Table 9 that compares IPv6 and IPv4 performance for ASes in SP and located at different hop-counts from their vantage point. The table shows that the similarity between IPv6 and IPv4 performance also holds at the finer granularity of paths with different AS hop-counts. Note that as mentioned earlier and unlike Table 7, because tunnels are unlikely to be involved, the definition of hop-count is in this case the same in IPv6 and IPv4.

	Penn	LU	UPCB
IPv6≈IPv4	92.3%	85.7%	72.2%
Other	7.7%	14.3%	27.8%
# ASes	13	42	36
x-check (+)	8	17	13

Table 10: World IPv6 Day – IPv6 vs. IPv4 for SP ASes.

In spite of the strength of the data presented in Tables 8 and 9, there are two caveats worth highlighting.

The first is that, as mentioned earlier, our coverage of the Internet is limited to a total of 1208/1785 ASes in IPv6/IPv4 (see Table 2) out of over 37,000 ASes in today's Internet. This is unfortunately not something we can address until more websites decide to become IPv6 accessible. The monitoring will be ongoing and we will provide updates in case new data invalidate earlier findings. The second caveat is that by all accounts the IPv6 traffic load remains light. As a result, one could argue that this low load does not stress potential forwarding bottlenecks. Equipment vendors report that IPv6 and IPv4 performance is now comparable, but direct observations consistent with our measurements would obviously be preferable. Ideally, this would consist of traffic measurement data for ASes on all monitored paths. Unfortunately, we did not have access to such data. However, we were able to carry out a (limited) experiment that paralleled our ongoing monitoring, but that took place during a period of time when IPv6 traffic was arguably higher.

Specifically, we configured our monitors to run monitoring rounds every 30mins during World IPv6 Day to the subset of sites that advertised participation in that event. As reported on the World IPv6 site itself (see http://www.worldipv6day.org/participants-dashboard), IPv6 traffic did spike during the event. Hence, it can be argued that limitations in IPv6 forwarding should have become more evident during that period.

The results are shown in Table 10 for sites/ASes in the SP category among the World IPv6 Day participants, and for all the vantage-points of Table 8 except Comcast for which the data was not available. As can be seen, the results are consistent and even to some extent better than those of Table 8, which further supports **H1**. Note the "expected" absence of a zero-mode row in Table 10 (all participants likely made sure that their end-systems were fully IPv6 qualified).

It should also be noted that while the outcome of the World IPv6 Day, as reported by its participants, is consistent with our findings, *i.e.,* IPv6 performance is on par with that of IPv4, it offered only limited conclusions. In particular, it relied on a small number of selected sites, which resulted in much sparser coverage, as seen when comparing the number of ASes in Tables 10 and 2. More importantly, it did not shed any light on hypothesis **H2** that identifies differences in routing choices as a major cause of poorer IPv6 performance. Validating **H2** is the topic of the next section.

5.4 Different IPv6 and IPv4 Paths

The next data sets we analyze are those associated with sites in DP, *i.e.,* sites in a destination AS reachable over different paths in IPv6 and IPv4. Recall from Section 4 that those sites (and ASes) will be used to

	Penn	Comcast	LU	UPCB
IPv6≈IPv4	3%	11%	10%	8%
Zero mode	12%	5%	3%	6%
# ASes	587	266	341	422

Table 11: IPv6 vs. IPv4 for DP destination ASes.

assess the validity of **H2**, namely, that differences in routing are the primary contributors to poorer IPv6 performance. As mentioned earlier, given that differences in routing (paths) is what primarily differentiates sites in DP from those in SP, attributing poorer IPv6 performance to routing is a natural thing to do. Table 11 reports the performance differences between IPv6 and IPv4 for ASes in DP. When comparing the results of Table 11 to those of Table 8, we see a clear difference. In particular, even when adding the set of ASes with a zero-mode to those for which IPv6 and IPv4 performed comparably (within our confidence interval), the fraction of ASes that see similar performance in IPv6 and IPv4 is much smaller than for ASes in SP. Note that unlike Table 8, Table 11 does not include any "x-check" rows. This is because deviations between IPv6 and IPv4 AS paths varied from vantage point to vantage point, which precluded meaningful comparisons.

Performance differences were also checked using data gathered during World IPv6 Day for sites in its DP set. Those results are presented in Table 12 that should be compared to Table 10. It again highlights a non-trivial difference between ASes in SP and DP, while also pointing to a substantially larger number of ASes in DP. These findings provide reasonable support for **H2**. Comparing Tables 11 and 12, we also note that World IPv6 Day participants fared better. This is not unexpected, since unlike Table 12, Table 11 likely includes sites with sub-par server IPv6 performance.

	Penn	LU	UPCB
IPv6≈IPv4	53.5%	48.9%	51.0%
# ASes	114	92	102

Table 12: World IPv6 Day – IPv6 vs. IPv4 for DP ASes.

As discussed in Section 4, data from other vantage-points was also used to asses the extent to which the network could be ruled out as the cause for performance differences. This called for identifying how often ASes in the IPv6 paths of ASes in DP are present in "good" IPv6 paths to ASes in SP. Table 13 reports the result of this investigation. It unfortunately shows that while most paths include a majority of known good IPv6 ASes, very few consist entirely of such good ASes. This does not imply that the other ASes were necessarily "bad," but only that they were not present in (good) AS paths to destinations in SP from other vantage-points. The presence of "bad apples" that contribute to the poorer observed IPv6 performance remains a possibility, although processing all ASes involved in monitored AS paths and looking for ASes whose presence would be associated with a high probability of poor IPv6 performance, did not reveal any such AS.

% good ASes in path	Penn	Comcast	LU	UPCB
100%	3.2 %	11.1%	6.4%	17.2%
[75% , 100%)	20.8%	8.3%	0.9%	22.4%
[50% , 75%)	58.8%	45.8%	68.8%	52.6%
[25% , 50%)	15.8%	27.8%	19.3%	7.8%
[0% , 25%)	1.4%	6.9%	4.6%	0%

Table 13: "Good" AS coverage in DP Paths.

5.5 Miscellaneous Findings and Limitations

The last "finding" we report on is a negative one. In particular, a question of obvious interest is whether sites/ASes that exhibit better IPv6 performance than IPv4 share some common property. We performed such an investigation, seeking to determine if such sites and ASes were more often associated with the DL, SP, or DP categories, or even if they were more frequently found in certain geographic areas. Unfortunately, no such grouping emerged, so that no dominant trait could be associated with better IPv6 performers.

Finally, although the study highlights IPv6 progress (hypothesis **H1**) and identifies steps to make it fully on par with IPv4 (hypothesis **H2**), it has limitations.

First and foremost is coverage. This could be improved by increasing the number of vantage points to provide greater geographic coverage and wider diversity in the type of network connectivity. Second, even if the analysis of SP paths is unlikely to be affected by tunnels, a more systematic investigation of their prevalence and impact would be desirable. In particular, this would allow more accurate comparisons of IPv6 and IPv4 paths of "equal" hop counts, as attempted in Table 7. A third limitation is that while we were able to obtain BGP routing tables after each monitoring round, a better option would be to retrieve path information at the same time as a site is being monitored, *e.g.,* by querying a local LookingGlass server. This is unfortunately difficult to realize across vantage points.

Another limitation of the work, one that is, however, not associated with its methodology, is the currently limited public access to its data, which would obviously be required to allow independent validation of the findings. Only a limited set of the data is currently accessible through the monitor's web front-end at http://mnlab-ipv6.seas.upenn.edu. We hope to soon be able to offer a `mysql` web interface that will support

queries to the underlying database. Additionally, we plan to make the full data sets available either through direct download or on a public repository, *e.g.,* such as Google's BigQuery.

6. CONCLUSIONS AND EXTENSIONS

Although adoption of IPv6 in today's IPv4 Internet remains nascent, there are signs of an improving situation; one that may ultimately avoid an unnecessarily drawn-out transition. The fact that IPv6 data plane support has improved and is now on par with IPv4 (hypothesis **H1**) is obviously positive, even if as mentioned in Section 5, it is not based on a "complete" sampling of today's Internet, *i.e.,* pockets of older devices with poor IPv6 performance are likely to remain. Additionally, the one factor that seems to bear the most responsibility for IPv6 performance lagging behind that of IPv4, *i.e.,* the use of less efficient paths (hypothesis **H2**), is one that can hopefully be remedied through focused but relatively standard measures, namely, peering parity between IPv6 and IPv4. In other words, the single most effective way to put IPv6 and IPv4 on an equal footing may well be to ensure peering parity.

There is obviously much work that can and should be done to confirm both hypotheses **H1** and **H2**, as well as understand how to best foster a rapid transition to IPv6. We have already mentioned the uncertainty related to the impact of growing IPv6 traffic on hypothesis **H1**. This calls for continuous monitoring to assess the extent to which a higher IPv6 traffic load becomes an issue, but the results from the World IPv6 Day seem encouraging. On a more pragmatic front, the ability of the monitoring tool and its underlying database to handle growth in IPv6 accessible sites beyond a few percents is unclear. This is something we will focus on to ensure that monitoring can continue at least until IPv6 adoption is reasonably far along. Finally and as mentioned earlier, in spite of the reasonable diversity of vantage-points we have achieved to-date, adding new ones to realize a more complete coverage of the Internet is important, if only to ensure that large pockets of poor IPv6 performance don't remain hidden.

Acknowledgments

The authors would like to thank their CoNEXT reviewers and their CoNEXT shepherd, David Malone, for inputs that have helped improve the paper. The authors are also greatly indebted to the many people who have helped with the deployment of the monitoring software at different vantage-points. Without their help and patience, in particular with early bugs, we would not have had access to all the data we were able to rely on. Special thanks are also due for the help we received in getting access to AS_PATH data from routers close to monitoring vantage-points. In particular, we would like to thank C. Chung (Comcast), M. Eagles (UPCB), S. Huque (Penn), and O. Maennel (Loughborough U.). Last but not least, P. Yan who developed the initial version of the monitoring software deserves credit for getting our early measurements off the ground.

7. REFERENCES

[1] Tracking the IPv6 migration. Arbor Networks Technical Report, 2010. Available at http://www.arbornetworks.com/research.html.

[2] K. Cho, M. Luckie, and B. Huffaker. Identifying IPv6 network problems in the dual-stack world. In *Proc. ACM SIGCOMM workshop on network troubleshooting: research, theory and operations practice meet malfunctioning reality*, 2004.

[3] L. Colitti, S. H. Gunderson, E. Kline, and T. Refice. Evaluating IPv6 adoption in the Internet. In *Proc. PAM*, Zurich, Switzerland, April 2010.

[4] A. Durand, R. Droms, J. Woodyatt, and Y. Lee. Dual-stack lite broadband deployments post IPv4 exhaustion. Internet Draft (Work in progress) (draft-ietf-softwire-dual-stack-lite-10), May 2011.

[5] R. Guérin and K. Hosanagar. Fostering IPv6 migration through network quality differentials. *ACM SIGCOMM Computer Communication Review*, 40(3), July 2010.

[6] E. Karpilovsky, A. Gerber, D. Pei, J. Rexford, and A. Shaikh. Quantifying the extent of IPv6 deployment. In *Proc. PAM*, Seoul, Korea, April 2009.

[7] Z. M. Mao, J. Rexford, J. Wang, and R. H. Katz. Towards an accurate AS-level traceroute tool. In *Proc. ACM SIGCOMM*, Karlsruhe, Germany, August 2003.

[8] S. Narayan, S. Kolahi, Y. Sunarto, D. Nguyen, and P. Mani. Performance comparison of IPv4 and IPv6 on various Windows operating systems. In *Proc. 11th International Conference on Computer and Information Technology (ICCIT 2008)*, December 2008.

[9] S. Narayan, P. Shang, and N. Fan. Performance evaluation of IPv4 and IPv6 on Windows Vista and Linux Ubuntu. In *Proc. 2009 International Conference on Networks Security, Wireless Communications and Trusted Computing*, 2009.

[10] Y. Wang, S. Ye, and X. Li. Understanding current IPv6 performance: a measurement study. In *Proc. 10th IEEE Symposium on Computers and Communications (ISCC 2005)*, June 2005.

[11] X. Zhou and P. Van Mieghem. Hopcount and E2E delay: IPv6 versus IPv4. In *Proc. PAM*, Boston, MA, April 2005.