Consolidated Review of
Internet Nameserver IPv4 and IPv6 Address Relationships

1. Strengths:

The paper presents analysis of server-side IPv6 deployment and provides a new and interesting perspective of the penetration and dynamics of server-side IPv6 addresses.

Two new measurement methods, plus discussion of what the data shows about DNS deployment.

The Akamai dataset provides broad coverage. According to the paper, some of the techniques are in production use at Akamai.

The techniques to form associations are neat (novel and non-trivial). The paper highlights the prevalence of different types of associations between IPv6 and IPv4 addresses. It presents methodologies to reduce complexity and understand the associations based on aggregation and temporal characteristics.

2. Weaknesses

The paper is poorly motivated, in that the work presented doesn't seem to address the motivations laid out [this issue was addressed in the camera ready revisions]. Although a stated motivation is geolocation, the paper analysis does not discuss IPv6 geolocation. It would be much more interesting if some of that work was included. While the paper highlights several interesting relationships between IPv4 and IPv6 addresses, it does not provide any indication of how useful these associations are in meeting the listed motivations for this work. The paper briefly discusses the usefulness for the associations for co-located services.

The technique can only be applied to name servers; it would be nice to be able to identify that a given router has both IPv4 and IPv6 addresses. The effect on geolocation accuracy was not assessed.

Outside of Akamai (and some similar CDNs), it's not clear who would be interested in this. Even many content providers like Google/Facebook do redirections based on client latency, not resolver location (which seems like a consequence of early Akamai design decisions rather than anything fundamental). While the data is interesting, the methods would be hard to replicate in a non-Akamai setting

The paper seems very long for the amount of insight gained. Could this have been a punchier short paper, or a long paper that included some of the uses of the data?

The paper's work doesn't have enough context with related work.

3. Comments

This is a very nice paper. The techniques are innovative, and the results are useful. This paper presented a nice discussion about evaluating DNS servers in both IPv4 and IPv6, including discussion of the implications of the measurement on DNS deployment. The approach is new and the discussion is carefully done.

The text at the top of page 2 repeats the abstract. Instead, it should explain it. In particular, the phrase "when the IPv6s are aggregated to /64 prefixes" is not an ideal way to explain that there is a one-to-one association between IPv6 /64 prefixes and IPv4 addresses.

I appreciate the care in discussing the techniques, although section 4 felt a bit verbose. It would have been nice to close the loop on the motivation about geolocation and discuss how the results affect IPv6 geolocation. A minor comment: in figure 5, the colors of the dots are hard to read. Two suggestions: replace the dots with squares so there is less whitespace (make the (x,y)=(2^0 , 2^0) point cover the range (x,y)=([2^0,2^1), [2^0,2,1)). Also label the lower left corner, where the dots are relatively sparse, with actual percentages. (Since that's the only place where they are not all <0.1%.)

I ended up feeling like the paper doesn't really deliver on the motivations / conclusions. Taking the motivations as laid out in the conclusion (similar in the intro):

i) Tracking the evolution of IPv6. It's not clear that tracking DNS pairs is a meaningful measure of evolution, and this paper doesn't concern itself with evolution.

ii) Understanding the potential for correlated failures and security risks when IPv4 and IPv6 services are physically or logically co- located. This paper doesn't address collocation (your tech report starts to address it), and it only looks at DNS

iii) Leveraging prior information of IPv4 addresses with respect to reputation and geolocation of newly active IPv6 addresses; This seems like an Akamai-specific motivation, and the paper doesn't really get into whether (even for 1:1 eq. classes) we actually know they are collocated, or what benefit we get from leveraging the info. I assume the techniques in this paper just generate pairs that are subject to further testing.

iv) Preventing erroneous Internet measurements intending to compare the performance of IPv4 and IPv6 paths. DNS doesn't seem like a common way to compare path performance, and your techniques are DNS-specific. Plus, the paper doesn't get into figuring out which v4/v6 addresses are not collocated.

Conclusion: "this paper is the first to take a comprehensive look at the server-side." I don't think looking at v4/v6 DNS pairs corresponds to a comprehensive look at server-side v6. DNS resolvers seem a small part of servers, and pairs seem only one thing to look at.

It seems like working through the collocation work in your tech report, plus perhaps adding a longitudinal study that demonstrates some interesting evolution, and ideally expanding beyond just DNS, would make this a much stronger paper.

Morley Mao's work on finding the association between clients and resolvers seems related, in that she also passed unique hostnames to clients to log which resolver ended up returning them. Your technique seems like a variation / new use of this previous work. In general, it's surprising not to find a related work section. One thing I was wondering about a comparison to is the IPv6 measurements that Lorenzo Colitti at Google did. Are their DNS techniques similar at all? Are there other similar works?

2.1.1: You mention that the load balancing can be advantageous for geolocation, but how do you know that all the servers are collocated?

Pg. 5: How do you know that NAT forwarders are not v6 enabled? I couldn't tell if you mean that you observed this in the anomalous measurements, if you mean it is common knowledge, or something else.

I'm very confused as to why your open resolver data set is so small, as I've seen a set that was nearly two orders of magnitude larger. While I don't know the details on how it was gathered, I believe it was also passively gathered (like yours). It seems surprising that anything would observe so many more than Akamai. Did you limit or filter your dataset in any way?

It's not clear what we're supposed to take away from the number of v6 addresses (say) in an equivalence class. Why is this important?

I'm curious about the prefix-level (at least for v6) results for "Restrict to final week."

You say that GoogleDNS may use some servers in local ISPs. Is it incorrect to assume this list is complete: https://developers.google.com/speed/public-dns/faq#locations ?

Table 3, last row: Actually, for this particular experiment, removing random pairs seems to have a big effect, more than just removing Google or just removed 6to4. The number of addresses after removing random pairs is much closer to the numbers after removing selected pairs than they are to the original numbers, although there are more pairs.

In Section 4.1.1, the restriction to shorter intervals increases the fraction of 1:1 correspondences due to reduced reconfigurations. Is it possible to extrapolate your results over multiple short instances, to obtain the "true" fraction of 1:1 correspondences?

Although your interest is in geolocation of name servers, it would be nice to be able to obtain these correspondences for routers as well, to allow comparisons between the IPv4 and IPv6 router-level topologies, or for servers, to assess the degree to which the IPv6 infrastructure is reusing IPv4 infrastructure.

Section 4.1.3: "8037-9582": remind the reader that this is "IPv4-IPv6"

4.2: I'm not sure what I was supposed to take away from this section. One thing seemed to be that you shouldn't have stopped at 200 probes.

4.3: It seems like the major takeaway is that, to get reasonable pairings, I should use a short dataset, exclude pairs that don't occur frequently, exclude 6to4, and exclude Google. Most of those seem like reasonable results (perhaps even ones you might have expected), but it took 12 pgs. to get here.

The section on aggregation to prefixes and ASes is quite long, and I missed what the actual implications are. Please make it more explicit.

Figure 5 is very hard to read in black-and-white. Instead, perhaps use a color scale such that the palest color is at one extreme, and/or vary the size of the dots based on the percentage.

To me, the first measurement technique is not "passive", since your measurement infrastructure sends packets.

The one day data set is rarely referenced in the paper. What additional insight did it provide?

Table 4: the paper mentions, the fifth row subsets have the same number of pairs as the fourth row, which only has 34 pairs. The table reports 11K+ pairs in the fifth row.

4. Summary from PC Discussion

This paper was discussed in the PC meeting. Most issues are already discussed in the reviews, but the PC would like to encourage the authors to make sure the motivation in the introduction aligns with the results in the rest of the paper, and if the data does not fully cover the stated motivation, the paper should clearly identify unsupported areas as future work.

5. Authors' Response

A major concern of the reviewers and the program committee is that "the work presented doesn't seem to address the motivations laid out." In the original version we used the term "motivations" in the loose sense of why someone might be interested in the work, and did not mean to imply that they were covered in presented results. In the revised paper, we restrict the motivation to what is explored in the body of the paper, which is our interest in associating Internet DNS resolver IPv6 address(es) with IPv4 address(es). The prior motivations are now labeled "potential future applications," and we clarify what aspects are contained in the paper.

Regarding GoogleDNS and the question: "Is it incorrect to assume this list is complete: https://developers.google.com/speed/public-dns/faq#locations," it may be of general interest to note that: yes, this list is incomplete. Our authoritative namesevers see queries from many more addresses in Google's AS 15169 than are listed on that web page.

In the revised paper we have attempted to address the comments of the reviewers. We would like to highlight here two important points on which the original text was unclear.

(1) Regarding the comment "the methods would be hard to replicate in a non-Akamai setting" - Actually, the technique can be implemented on any authoritative nameservers under common control along a DNS namespace hierarchy.

(2) Regarding the comment that the first measurement technique is not passive, since the measurement infrastructure sends packets - We consider the technique to be passive as the collection of address pairs is simply a by-product of the client's nameserver resolving a domain under our authority that the client requested. No additional packets are sent when one deploys the technique for domain names that are being used anyway.

We want to thank the reviewers for their time and careful reading our paper, and their many constructive suggestions.