Consolidated Review of
Exploring EDNS-Client-Subnet Adopters in Your Free Time

1. Strengths

This paper explores how we can use the EDNS-Client-Subnet DNS extension to map out CDNs and content providers (CPs) that support it (currently only a small number, such as Edgecast and Google). This extension carries along the client's prefix with a DNS request, allowing DNS redirection based on client address rather than resolver address. Because the requests are not authenticated in any way, the authors can query a CP with an arbitrary (or all) prefixes in order to understand where the CP would map the client. Because Google is the largest adopter, many of the measurement results focus on Google.

Given how little information CDNs/CPs sometimes expose, it's great to take advantage of opportunities like Client-Subnet presents. Interesting results on the mapping granularities that Google uses. The paper was generally well-written and easy to follow.

The methodology presented is novel. The experiments undertaken and reported on are not necessarily deep, but do well illustrate the technique and show its promise. The execution of the experiments is sound.

The paper offers an interesting first look at how EDNS can be used to infer various properties of the network, including user-to-server mappings, EDNS deployment, and so forth.

2. Weaknesses

The work seems premature, even for a short paper. Asks interesting questions, but leaves many of them (and unasked ones) open.

By not deaggregating prefixes whose responses are scoped to something finer-grained, the methodology may not expose the entire footprint of a provider, and questions like whether or not an entire AS is sent to the same servers cannot be definitively answered.

While the framing leads one to believe the technique will illuminate the client-to-server assignment behavior, most of the paper is actually about the workings of the ECS mechanism itself. I think this could be better framed a bit. I.e., the paper shows the efficacy of the technique and a short case study that offers an initial treatment of Google's assignment scheme.

The paper presents a nice idea with some early results that show the power of the approach. However, the paper seems in a very early stage, giving the impression that a more exhaustive and deeper analysis can be performed using this measurement approach (e.g., especially regarding the user-to-server mapping).

The approach is somewhat obvious and unsurprising, the results are specific to the particular dataset (with no attempt to interpret or generalize), and the tone of the paper is overblown. Some important validation is missing.

3. Comments

It's neat to try to understand how these large providers operate. However, I think it's worth pushing further to understand more before publishing this work. On its own, the observation that one can use ECS in this way is a very minor contribution. I think most people in the measurement community who are familiar with the extension will recognize this potential (from reading the Faster Internet page, where the dig patch performs the measurement you describe), and the extension has been publicized in the community via the NWU papers that you cite. So, I'd really like to see the paper glean more insight from its use. The paper does some of that, but leaves many questions open.

Nice to see the data and tooling will be made available to the community.

Yes, the IP address/prefix that is issuing the DNS query is useful, and being able to set this value can proxy for actually performing the measurements from different vantage points (presuming that the DNS resolvers cannot determine that the queries are, in fact, coming from a single location). But, the last point is not validated, and it would have been easy to validate by performing EDNS from several locations, and swapping IP addresses at each location to verify that, in fact, the resolver was only using the EDNS value to perform the resolution. The paper claims: "In principle, if the prefix length corresponds to a publicly announced prefix, one may expect that the returned scope is equal to the prefix length." I could imagine many reasons why this would not be the case (anycast prefixes, provider prefixes, etc.) This seems like a naive hypothesis, and the surprise at the result feels forced.

A number of issues could use further exploration, and I'd be interested to see what you find:

- The extension tries to make the Internet faster by making it easier to direct clients to nearby servers, but it potentially makes things slower by reducing DNS cache hits and making DNS slower (each result applies to fewer users, causing fewer users to share hits, and there are more possible cache entries, leading to earlier eviction). Can you use data similar to what you used to generate the PRES dataset and the data from 5.2 to try to study this tradeoff?
- In 3.2 the procedure for vetting ECS adopters isn't great. It took several passes before I was mostly sure I understood what you were talking about. (And, I am not entirely sure now.) You might try to clean this up and make things more concrete. Perhaps run through an example.
- 3.2: "The second group, about 10%, seems to be ECS-enabled but does not appear to use it for the tested domains." <-- I am just not sure I follow what you're saying here. If ECS isn't used for the tested domains then what is the basis of the guess ("seems to") that these are ECS enabled. I just don't get it.
- Section 4: "We emphasize that a single vantage point is sufficient for performing our experiments." Do you know this? It certainly follows intuition, but the strength of the statement would be a bunch higher if you actually tested it. It is conceivably possible that both the advertised prefix in a query and the source IP address of that query are both serving as input, right? Why? I have no idea. But, at least a sanity check from another location would seem useful here.
- 5.2 has some interesting results that I hope you explore further. When the scope/answer changes quickly, is it switching between Google and GGC? Are the multiple answers near to each other? What do you think is going on? Similarly, you

mention looking further into Google's scoping at /32 for CDN servers.

- 5.3: neat examination of consistency of mapping across an AS. When an AS is served from multiple /24s, are the /24s generally all within one AS? Are the ASes that map to a single /24 usually ones that announce only a little prefix space, whereas ones mapped to multiple are those that announce more address space? You mention that Google maps a small number of ASes to a large number of server /24s, and that many of these ASes have large footprints. Do some of them not have large footprints? What is going on in those cases? For the /24s that serve only a single AS, you say that they are typically GGC. Are they always GGC? If not, what are the other examples?
- 5.3: Please run your churn experiment for longer, seems interesting. Curious why you think we'd expect higher churn. I would think that Google would try to have enough capacity to always serve from the lowest latency site, and so load balancing and churn would be minimized. Intro claims that the underlying assumption of standard DNS resolution is that the end-user is close to the local resolver. However, there's no reason for services that operate their own DNS and their own servers to need this assumption. [19] shows a simple mechanism to figure out which clients use which resolvers, Facebook blogged a few years ago about the same idea, and [17] seems to state that redirection is based on latency measurements to clients. So, in that situation, the problem is not that clients may not be near their resolver (on its own, that is surmountable), but rather that a resolver may serve clients that are not near each other.
- Is there reason to believe that /24 is the granularity at which Google divides their servers internally? The paper seems to assume that (primarily in 5.3), but I wasn't sure whether or not it was supported.
- Given how quickly you can query for even your largest prefix sets, I didn't find the comparison of different sets of network prefixes interesting. Why wouldn't one just use all routable prefixes? And, why wouldn't one (de)aggregate them based on whatever ECS scope is returned (5.2 suggests we should at least de-aggregate to /24 for Google)? The one part of the prefix set comparisons that I found interesting was that "prefixes served by the neighbor ISP are from a customer." I would think that the likely explanation is that the ISP tells Google which blocks it is willing to serve, which isn't quite Google inferring it (which you claim). Based on what you know about address allocation, do the prefixes in the PRES datasets likely host clients or just resolvers and other infrastructure?
- Similarly, if Google is using a finer-grained scoped than the BGP prefix, it would be interesting to see if you can find differences between those subprefixes, to try to understand why Google is making this decision: are they located in different areas? Do traceroutes to them look different? RTTs?
- 5.1: The GGC data is interesting! You talk about how many server IPs are outside Google's ASes. How many clients are directed to these servers? In other words, even though most of the IPs are caches, could they be small caches that each only serve a small number of client prefixes (you have results in 5.3 showing that is sometimes the case)? I think we would expect this if they are low in the AS hierarchy, right?
- In 5.2 I was wondering whether the TTLs varied between ECS and non-ECS replies. Given that ECS replies can be quite specific (which you should when the scope is assigned to /32s pretty often), can the mapping be longer lived? I am not sure I have a bunch of intuition here, but it seems useful to explore.
- Figures 2(a) and 2(d) are a little bit difficult to follow because the dots are on top of one another. Why not show something more direct, like error? Why are the scatter plots for the other comparisons so different? Are there any general conclusions here?
- All of the observations on Page 5 simply read off the results from the plots but do not offer any insight for why we are seeing specific results.
- The insights in Figure 3 are pretty meaningless because they are not analyzed by geography. For example, it is likely that in the US, different ASes have unique subnets, whereas in more rural or developing regions, assignment to subnets is sparser. The analysis here is pretty thin.
- 5.3: Neat graphs!
- I found the description of the prefix lists from various places as "datasets" to be sort of strange. Perhaps just style, but those lists seem more like, well, lists to me.

4. Summary from PC Discussion

This paper was discussed at length both in online discussion on the submission site and at the PC meeting itself. The PC generally found that the findings in the paper were enough to warrant contribution for a short paper, but several reviewers were somewhat disappointed with the overblown claims about the paper's contributions, which turn out to be far more modest than the paper's introduction suggests. For example, the measurement method was not invented by the authors, but is in fact well-described on the EDNS Google project site. The paper should offer more credit where credit is due. Even giving credit to other previous work, the paper represents a nice short paper in its own right. There is no reason to overstate claims.

Ultimately, the PC decided to take the paper because they felt it was a useful first step in studies of EDNS and others may decide to build on this study. That said, much of the discussion focused on the preliminary nature of the results---especially in comparison to the claims---which nearly caused the paper's demise. This is a useful short paper that should be published, but the paper could be improved with more discussion of how the work leads to avenues for future work.

5. Authors' Response

The paper provides the toolbox, traces, and preliminary analysis of the collected data to shed light on the deployment and operation of some CDNs that have adopted EDNS-client-subnet. This is important given the central role that such CDNs, e.g., Google, play in today's Internet.

Claims: The paper demonstrates that it is easy to take advantage of the adoption of EDNS-client-subnet to (i) uncover the global footprint of such adopters, (ii) infer the DNS cacheability of such adopters for any arbitrary network, and (iii) to capture snapshots of the user to server mapping as practiced by the adopters using a single vantage point. Previous attempts to measure large CDNs, e.g.,

Google, required well distributed vantage points that are typically difficult to have access to. We admit that the claims were misleading in the submitted version and we fixed that in the camera ready.

Method: Our method was criticized because it did not used de-aggregated prefixes. In the revised version we make clear that public prefix sets, i.e., RIPE and Routeviews have significant overlap of prefixes. This is due to the fact that the network announcements are collected at different networks in terms of size and Internet hierarchy. For example, in the RIPE prefix, if we consider the less specific prefix (excluding private and non-valid prefixes) we end up with about 131K prefixes. Thus, when we utilize the RIPE prefix set we use de-aggregated prefixes (many of length /24 or more specific) of the above-mentioned 131K non-overlapping ones. To de-aggregate the RIPE (or Routeviews) prefix set to /24 prefixes is not desirable as it is not easy to validate which /24s contain active IPs and also significantly increases the running time of the experiment, at least when a single vantage point is utilized. Nevertheless, we confirmed that the set (number) of Google server IPs that our method uncovers differs by less than 5% (1%) from the set that a /24 prefix de-aggregation method uncovers [see Calder et al. IMC 2013]. Thus, the results of the two methods converge to similar sets and numbers of Google server IPs but when RIPE (or Routeviews) prefix set is used the number of queries sent is way smaller.

Credit: We regret if we failed to give credit to the developers of tools and software we build upon. We did our best to include all the related references in the camera ready including links to blogs of operational community that have first identified some of the shortcomings of the adoption of EDNS-client-subnet.

Vantage Point: We confirmed that by simultaneously utilizing different vantage points, namely, two vantage points in other residential networks and another one in a University in the US, the results were almost identical. It is also easy to scale up the query rate by using multiple vantage points in parallel, e.g., by utilizing PlanetLab nodes, but our experiment demonstrates that a single vantage point is good enough for the purpose of this study. As we use a single vantage point in our experiment any end-to-end measurement including traceroutes, DNS overhead, etc. is out of the scope of this study.

Datasets: In our study we use a number of datasets that we also make publicly available. We use different private and public prefix sets to query a number of EDNS-client-subnet adopters in an attempt to find the best prefix set to uncover the CDN footprint and capture snapshots of the user to server assignment as performed by the CDN. In the camera ready we also present a newset of traces that span more than four months to track the expansion of CDN footprints.

Analysis: We provide insights regarding the analysis of the collected data in the camera ready. For example, we provide information about the networks that host Google servers, we analyze consecutive snapshots of assignments of users to Google servers during a period of 48 hours, and analyze additional datasets that we collected after the submission of the paper to confirm some of our initial observations.

Future work: This paper takes a first step to uncover details about EDNS-client-subnet adopters deployment and operational practices. Given the fact that some important CDNs such as Google and Edgecast have already adopted this DNS extension, and a number of ISPs and CDNs are considering adopting it as well, it is useful to raise awareness of some of the consequences in a systematic way. The tools and datasets produced in this work can be used by researchers in a number of ways. The most obvious one is to study the expansion of CDN footprints in different networks and locations over time. This becomes increasingly important as many CDNs continuously deploy servers at the edges and is not anymore enough to rely on the AS number to infer the organization of an IP. It is also possible to study the dynamics of user to server assignment over time especially when with information about outages, flashcrowds or other events. EDNS-client-subnet may use this tool to evaluate what type of information can be inferred with this technique and come up with countermeasures for better protect sensitive business information.