

If a Path is Inflated, and Noone Uses It, Is It Inefficient[a]?

Anonymized

ABSTRACT

Anycast is a means of distributing content that has been praised for its simplicity and performance, yet criticized for inflating user latencies in some cases. However, anycast CDNs and anycast DNS resolvers serve latency-sensitive content to millions presenting an apparent contradiction[b][c][d]. To resolve this discrepancy in understanding, we analyze anycast in two important settings: the root DNS and a large anycast CDN, placing an emphasis on how anycast’s performance translates to end user latency. First, we question whether latency matters at all, and find that while it does in the CDN setting, it makes little difference in the root DNS due to heavy caching near end users – root DNS resolution only accounts for at most 2 ms per day for most users. We then examine how anycast specifically impacts performance to end users, and again demonstrate users see negligible effects from anycast path inflation. Users of a large anycast[e][f][g] CDN, conversely, can experience a few ms per page load of anycast path inflation, with the effect becoming more prevalent in larger anycast deployments. The effects of the inflation, however, are dwarfed by the performance benefits users see by, on average, visiting closer anycast sites[h][i]. Hence we demonstrate that although anycast can inflate latencies, the degree to which this affects users largely depends on the setting, and the prevalence of this inflation does not tell the whole story[j] – anycast still does a good job routing users to sites.

1 INTRODUCTION

IP anycast, an approach to routing in which geographically diverse servers known as anycast replicas all use the same IP address, is used by a number of operational DNS [?] and CDN [?] systems today, in part because of its support to improve latency to clients and decrease load on each anycast server [?]. However, studies ([?]) have argued that anycast provides sub-optimal performance for some users, compared to the lowest latency one could achieve given deployed replicas. Despite results demonstrating theoretical routing inefficiency, how these inefficiencies impact user experience is not well understood.[k] Anycast, therefore, presents a paradox providing benefits of increased capacity and

decreased latency while also purportedly hurting performance.

To resolve this, we take a step back and ask pointed questions – why do distributed systems such as CDNs and operational DNS use anycast, how does anycast affect their performance, and how does this performance ultimately translate to end user experience? Towards answering these questions we analyze anycast from two different angles: the root DNS and a large anycast CDN, chosen for their overlapping, yet distinctive, goals. The root DNS servers feature in studies [?] involving anycast because it is relatively easy to gain access to root DNS data, because it is straightforward to gain access to information about their deployments, and because they are run by several organizations [?]. This last fact manifests itself in a diverse set of deployment strategies, for essentially the same service. Here, examining the root DNS and anycast CDNs is particularly interesting because this analysis illustrates how the setting in which we study anycast heavily influences the conclusions we can draw – while we find that mitigating anycast path inflation is quite important for anycast CDNs, the impact of latency *at all* in the root DNS setting is negligible.

Although we do wish to examine how anycast affects performance both at the roots and in anycast CDNs, we first take a step back and examine whether performance (that is, latency) matters *at all* (section 3). By leveraging global root DNS traces and user data from a large anycast CDN, we demonstrate that the effect of root DNS latency on user-perceived latency is negligible, accounting for perhaps a few milliseconds of wait time per day or fractions of a percent of a page load. This is due mostly to heavy caching of root DNS records; hence, regardless of how inflated paths to the root replicas are, this latency is amortized over large user populations. Conversely, we show that latency matters considerably for anycast CDNs, comprising tens of percents of page load time (section 5). Collectively these results suggest organizations managing the root servers have little incentive to improve latency and, consequently, it makes little sense to evaluate the latency achievable with anycast by looking at the root servers.

With these basic results about latency, we revisit frequently posed questions about how anycast specifically

impacts these services. We find that, even though round-trip times differ significantly by root DNS anycast deployment size, these differences are negligible when looked at from a per-page load perspective, making at most an `INSERT_NUMBER[l]` millisecond difference. Similarly, even though we find that increasing deployment size can lead to more inflation in the roots[m], this inflation negligibly factors into page load times (section 4). Conversely, we find that for an anycast CDN, although increasing deployment sizes does make anycast path inflation more prevalent, the latency per page load decreases by tens of milliseconds (section 6) with additional sites. Moreover, regardless of deployment size, the path inflation for the large anycast CDN is less than half that of the roots for more than 90% of users (of the large anycast CDN). Hence larger deployment sizes can provide tangible latency benefits to anycast CDNs, but probably provide little benefit in the root DNS setting, and the magnitude of these benefits are largely dependent on deployment details.

2 BACKGROUND: ANYCAST IN DISTRIBUTED SYSTEMS

IP anycast is a system in which geographically diverse servers known as anycast replicas all use the same IP address, and therefore rely on BGP’s notion of propagating the best route towards the destination. We discuss why distributed services on the Internet may wish to use Anycast, and ultimately leverage this knowledge in later sections to discuss why the root DNS and CDNs may use IP anycast. The benefits we are about to mention are *potential* – the realization of such benefits depends on the deployment details of the system using IP anycast. We also provide some background on the two services we investigate – the root DNS and a large anycast CDN, as this context will prove useful in later sections.

2.1 Potential Benefits of Anycast

[why distributed systems might use anycast, different performance benefits they might shoot for] IP anycast is first and foremost, simple and scalable. Network managers offload the responsibility of mapping users to sites to the network, and may add more sites by simply advertising the address from more locations. IP anycast can provide low latencies from users to destinations, since it relies on the network’s notion of “best path”; practically, this is handled by the Border Gateway Protocol (BGP). Anycast also offers resiliency in that, should one site go offline, the network will handle routing users to different sites that are still running. A final potential benefit offered by anycast is load balancing. The idea is

that users are spread out over the network, and so they will roughly be routed to corresponding anycast sites in different areas of the network [?]. Although recent research suggests anycast may not naturally balance load as was previously thought [?], what is more potentially useful is that anycast routes traffic predictably. Prior work suggests that only a very small fraction of anycast paths are unstable [?], and so network managers may provision in advance for expected (if unbalanced) load.

2.2 Root DNS anycast

[brief description of dns resolution process, caching potential] DNS is a fundamental service on the Internet that maps human readable hostnames to IP addresses [? ?]. To resolve a hostname, a user will send DNS requests in the form of UDP packets to one or more recursive resolvers (RRs) provided by their ISP¹. The RR then requests the records from a root DNS server, top level domain (TLD) server and authoritative DNS (ADNS) server corresponding to the record the user requested. Since each request is a correspondence between the RR and a remote server, there can be several requests made by the RR for a single end-user request.

The root DNS servers [?] are grouped into thirteen letters, and are managed by twelve distinct organizations. Each letter consists of a certain number of anycast replicas, with actual numbers ranging from a few to a few hundred, and each letter is assigned a unique v4 and v6 address. Each replica serves DNS records for the TLD records, of which there are a little over 1,000. All but two of these records has a TTL of two days, with the exceptions having TTL’s of one day. When a RR needs to query a root server, it may query whichever one it wishes (subject to network administration policy); however, recursive resolver software is known to query high performing (that is, low latency) letters more often. The performance of anycast in the root DNS has been extensively studied in a variety of contexts [? ? ? ? ?].

2.3 CDN Anycast

[description of microsoft’s CDN, with a focus on intricacies of rings, and how we use them to simulate deployments] To investigate anycast[n], To study anycast in a setting that contrasts the root DNS, we leverage a large anycast CDN that serves millions of users from more than 100 front ends. Traffic destined for the large anycast CDN will enter the CDNs network at a peering point, and be routed to one of the anycast replicas serving the

¹The user can specify whatever RR they wish, but one can sensibly assume the typical user’s RR is set by the ISP, broadcasted through DHCP.

If a Path is Inflated, and Noone Uses It, Is It Inefficient[a]?

content. This large anycast CDN has a logical hierarchy of layers, called rings, that serve different types of content. Hence, traffic from a user prefix destined for the CDN may end up at two different front ends (depending on the service), but will ingress into the network at the same peering point.

In the following, we use these rings to simulate anycast deployments of different sizes, as a way of coarsely assessing anycast's behavior with increasing deployment size. That is, since various services use different sized rings, we are able to isolate performance metrics for each ring. Although the front ends serving content are anycasted, this is not a perfect analogy to actually deploying different sized anycast networks on the Internet, since traffic traversing the public Internet will take the same path, regardless of the ring in question. Hence we are, to some extent, only assessing anycast within the CDNs network.

3 DOES DNS ROOT LATENCY MATTER?

Before answering questions about anycast latency, we would first like to understand root DNS latency. Here we not only measure what root DNS latencies are[o], but also how this ground truth latency manifests itself in *user perceived* latency. We approach this from two main perspectives: local and global. The former allows us to estimate the fraction of page load time (PLT) during which a client is waiting for root DNS resolution and the latter allows us to estimate the total time per day a user is waiting for root DNS resolution.

[ISI results – root dns latency cdf and PLT implications Main idea is cache hit rate is high, and root requests are rarely generated.]

[DITL – root DNS per day latency; Main idea is that amortizing requests seen to the roots over large user populations makes latency implications small]

[Embedded in the above is a comparison between high latency and low latency roots with implications to user latency, so No, latency doesn't matter]

4 ANYCAST PERFORMANCE IN THE ROOT DNS

[anycast path inflation per RTT & per page load; Main idea is path inflation does become more prevalent with increasing deployment, but makes no difference to users] [bar graph showing which sites are hit, corroborates the idea that inefficiency grows with deployment size, usually]

[use these results to suggest sites are added to root deployments for resilience, since there is no difference from a PPL perspective]

5 DOES ANYCAST CDN LATENCY MATTER?

[latency per RTT and per page load at various ring sizes; Main idea is RTT latency (and perceived latency) is significant from a PPL perspective, and ?depends on the application?] [compared to root DNS per page load; idea is CDN latency is orders of magnitude more significant PPL, and latency PPL 10th percentile vs 90th percentile is large]

6 ANYCAST PERFORMANCE IN CDNS[P]

[path inflation per RTT/page load, by ring; Main idea is that inflation becomes more prevalent, but latency PPL still goes down] [inefficiency by ring; shows fewer users go to the closest site, but still latency PPL goes down] [geographic path inflation per RTT/page load, compared to the roots; CDN inflation < root inflation,[q] argue that CDN works to control it via peering[r]] [case studies of intermetro variability, or unexpectedly poor performance, highlighting the intricacies (time permitting)]

7 RELATED WORK

IP anycast performance is usually studied in the context of two applications: the root DNS servers, and CDNs. In addition to these topics, we discuss studies of popular recursive resolvers, and user-centric measurements of web performance.

7.1 Root DNS Anycast Performance

The performance of anycast in the context of root DNS is generally gauged by anycast's ability to balance load among server replicas or provide low latency to users. Generally, all studies conclude that anycast successfully balances load, while latency performance depends on the specific deployment configuration. [?] looks at a DDoS attack on the root name server infrastructure, and generally shows that anycast is a good defense mechanism against such attacks. An earlier study, [?] confirms that anycast protects the root DNS infrastructure against such attacks and, furthermore, that anycast routes users to an optimal location in most cases. [?] looks at user latency to C, F, K, and L-root and attributes better performance to good geographic location and peering strategies. These findings coincide with an earlier study, [?], who conclude the performance of anycast is intrinsically linked to deployment strategy. Additionally [?]

finds that as few as 12 sites can provide “good” latency to users. [?], [?], [?], and [?] are all examples of studies who quantify latencies to various root servers, and note how these compare to the (optimal) latency of the closest unicast alternative for the user who issued the query.

7.2 CDN Anycast Performance

Some CDNs (e.g. Cloudflare, Edgecast, Fastly) use IP anycast to augment their serving infrastructure. When deploying an Anycast CDN (ACDN), delivering content to users with low latency becomes a high priority, as there is a large financial incentive to do so. The simplicity of IP anycast comes at the cost of having coarse grained control over where user queries land. Shifting user load between nodes during peak hours, for example, is a challenging problem. As a potential solution, [?] and [?] use DNS redirects at ADNS servers to shift load among anycast nodes, albeit in slightly different ways. [?] analyzes what latency users are achieving, compared to optimal, when being routed to anycast nodes and finds that 10% of users experience a latency inflation of at least 100 ms.

7.3 Recursive Resolvers and the Benefits of Caching

Similar to the RR analysis conducted here, [?] looks at DNS traffic on a small network and notably finds that 16% of queries resulted in queries to the root, most of which were for invalid domains. As this study is quite old, it is no surprise that this rate has decreased (recall we observed .5% of queries resulted in queries to the root) since browser designers and network engineers understand the importance of caching. [?] also looks at a RR and analyzes statistics of DNS exchanges occurring over it including DNS transaction latencies. Both [?] and [?] look at certain pathological behaviors of popular recursive resolvers, and the implications these behaviors have on root DNS load.

7.4 Web Performance

Although we were unable to find any specific study that looked at how web performance and root DNS latency were related, there are certainly studies characterizing web performance. [?] characterizes web performance bottlenecks in (at the time) new broadband networks, and finds that latency is the main bottleneck for PLT when the user’s bandwidth exceeds 16 Mbps. However, the study does not realistically emulate a page load and, in particular, can not analyze the effect of having multiple DNS resolutions per page. Similarly, [?] analyzes how each step of a page load contributes to the aggregate

PLT using a tool designed in-house. However, unlike [?], they did not conduct a large measurement campaign and do not include information about multiple DNS lookups per page. A more recent study, [?] provides a brief survey of web performance measurement studies and explains why it is difficult (with current practices) to compare two different studies in web performance.

8 CONCLUSION

[a]ideally we would control the linebreak in the latex so it occurs after “it,” [b]I’m not sure it is a contradiction. Aren’t the people criticizing it well aware that it serves many users? [c]I was trying to get across that it is odd for (for example) Microsoft to use anycast to serve latency sensitive content if it is known Anycast inflates latencies. [d]Yes, I think your instinct of wanting to convey that is good, I was just suggesting thinking a bit more about the phrasing. [e]What is “the” CDN? [f]I mean to change this to a macro, switching between “a large anycast CDN” and “Microsoft’s anycast CDN”.

Do you think I need to talk more about what sort of data I use to draw these conclusions in the abstract? [g]You might consider a brief phrase like “Analyzing global traces gathered from a large anycast CDN and from all root DNS servers worldwide”. Not necessary in the abstract, but can help. In the intro, you definitely want to talk a little about your data. [h]It isn’t clear what this is referring to, closer than what, etc [i]Since these results and what we want to say about them are in flux, I will wait before clarifying. [j]Will we also have an argument that it is less prevalent for CDNs? The analysis is in flux, and the data isn’t perfect. [k]This seems important but is buried in the middle of a paragraph [l]I recommend not putting any placeholders in without markup. We should use markup like **TBD: or** that makes it colored, bold, and perhaps renders as something like `[[[TODO: INSERT NUMBER]]]` [m]you use this phrase a lot in the abstract and intro, but I’m not sure it’s clear what you mean [n]+mcalders@microsoft.com Is this a valid description of Microsoft’s anycast network, in the context that we leverage it? Please let me know if you think any major details are missing. [o]perhaps a quick set of speed checker measurements [p]this analysis is in flux [q]might be an unfair comparison since the root DNS has many more nodes than Microsoft [r]and more active debugging of bad routes, probably [s]interesting that in 10 minutes they observe so many queries for COM TLD yet don’t see any issue with that [t]Might be an interesting tool to use [u]Mark shared in an email – shows time between DNS queries & TCP connection starts can be big, which suggests DNS is not blocking

REFERENCES

- [1] root-servers.org
- [2] akamai.com/us/en/multimedia/documents/white-paper/akamai-designing-dns-for-availability-and-resilience-against-ddos-attacks.pdf
- [3] dns-oarc.net/oarc/data/ditl/2018
- [4] The Top 500 Sites on the Web. alexa.com/topsites
- [5] M. Allman. On Eliminating Root Nameservers from the DNS. In *Proceedings of the 18th ACM Workshop on Hot Topics in Networks*,. 1–8 2019.
- [6] H. A. Alzoubi, S. Lee, M. Rabinovich, O. Spatscheck, and J. Van Der Merwe. 2011. A practical architecture for an anycast CDN. *ACM Transactions on the Web (TWEB)* 5, 4 (2011), 17.
- [7] Amazon. aws.amazon.com/route53/faqs/
- [8] Amazon. aws.amazon.com/cloudfront/
- [9] A. S. Asrese, P. Sarolahti, M. Boye, and J. Ott. WePR: a tool for automated web performance measurement. In *2016 IEEE Globecom Workshops (GC Wkshps)*,. IEEE, 1–6 2016.
- [10] H. Ballani, P. Francis, and S. Ratnasamy. A measurement-based deployment proposal for IP anycast. In *Proceedings of the 6th ACM SIGCOMM conference on Internet measurement*,. ACM, 231–244 2006.
- [11] M. Calder, A. Flavel, E. Katz-Bassett, R. Mahajan, and J. Padhye. Analyzing the Performance of an Anycast CDN. In *Proceedings of the 2015 Internet Measurement Conference*,. ACM, 531–537 2015.
- [12] T. Callahan, M. Allman, and M. Rabinovich. 2013. On modern DNS behavior and properties. *ACM SIGCOMM Computer Communication Review* 43, 3 (2013), 7–15.
- [13] Cloudflare. Load Balancing without Load Balancers. blog.cloudflare.com/cloudflares-architecture-eliminating-single-p/
- [14] Cloudflare. What is DNS? cloudflare.com/learning/dns/what-is-dns/
- [15] L. Colitti, E. Romijn, H. Uijterwaal, and A. Robachevsky. 2006. Evaluating the effects of anycast on DNS root name servers. *RIPE document RIPE-393* 6 (2006).
- [16] R. de Oliveira Schmidt, J. Heidemann, and J. H. Kuipers. Anycast Latency: How Many Sites Are Enough?. In *International Conference on Passive and Active Network Measurement*,. Springer, 188–200 2017.
- [17] T. Enghardt, T. Zinner, and A. Feldmann. Web performance pitfalls. In *International Conference on Passive and Active Network Measurement*,. Springer, 286–303 2019.
- [18] A. Flavel, P. Mani, D. Maltz, N. Holt, J. Liu, Y. Chen, and O. Surmachev. Fastroute: A scalable load-aware anycast routing architecture for modern CDNs. In *12th {USENIX} Symposium on Networked Systems Design and Implementation ({NSDI} 15)*,. 381–394 2015.
- [19] S. Goldlust. Early Refresh of Cache Records. kb.isc.org/docs/aa-01122
- [20] Google. developers.google.com/speed/public-dns
- [21] J. Jung, E. Sit, H. Balakrishnan, and R. Morris. 2002. DNS performance and the effectiveness of caching. *IEEE/ACM Transactions on networking* 10, 5 (2002), 589–603.
- [22] D. Katabi and J. Wroclawski. 2000. A framework for scalable global IP-anycast (GIA). *ACM SIGCOMM Computer Communication Review* 30, 4 (2000), 3–15.
- [23] J. Kurose and K. Ross. 2010. Computer networks: A top down approach featuring the internet. *Peorsoim Addison Wesley* (2010).
- [24] M. Lentz, D. Levin, J. Castonguay, N. Spring, and B. Bhattacharjee. D-mystifying the D-root Address Change. In *Proceedings of the 2013 conference on Internet measurement conference*,. ACM, 57–62 2013.
- [25] Z. Li, D. Levin, N. Spring, and B. Bhattacharjee. Internet anycast: performance, problems, & potential.. In *SIGCOMM*,. 59–73 2018.
- [26] J. Liang, J. Jiang, H. Duan, K. Li, and J. Wu. Measuring query latency of top level DNS servers. In *International Conference on Passive and Active Network Measurement*,. Springer, 145–154 2013.
- [27] S. McQuistin, S. P. Uppu, and M. Flores. Taming Anycast in the Wild Internet. In *Proceedings of the Internet Measurement Conference*,. 165–178 2019.
- [28] C. Metz. 2002. IP anycast point-to-(any) point communication. *IEEE Internet Computing* 6, 2 (2002), 94–98.
- [29] P. Mockapetris. Domain Names - Implementation and Specification. ietf.org/rfc/rfc1035.txt
- [30] G. Moura, R. d. O. Schmidt, J. Heidemann, W. B. de Vries, M. Muller, L. Wei, and C. Hesselman. Anycast vs. DDoS: Evaluating the November 2015 root DNS event. In *Proceedings of the 2016 Internet Measurement Conference*,. ACM, 255–270 2016.
- [31] C. Partridge, T. Mendez, and W. Milliken. Host Anycasting Service. tools.ietf.org/html/rfc1546
- [32] S. Sarat, V. Pappas, and A. Terzis. On the use of anycast in DNS. In *Proceedings of 15th International Conference on Computer Communications and Networks*,. IEEE, 71–78 2006.
- [33] S. Sundaresan, N. Magharei, N. Feamster, R. Teixeira, and S. Crawford. 2013. Web performance bottlenecks in broadband access networks. *ACM SIGMETRICS Performance Evaluation Review* 41, 1 (2013), 383–384.
- [34] Verizon. verizondigitalmedia.com/media-platform/delivery/network/
- [35] Y. Yu, D. Wessels, M. Larson, and L. Zhang. 2012. Authority server selection in DNS caching resolvers. *ACM SIGCOMM Computer Communication Review* 42, 2 (2012), 80–86.
- [36] zakird. Private IP List. github.com/zmap/zmap/blob/master/conf/blacklist.conf