Characterization of Anycast Adoption in the DNS Authoritative Infrastructure

Raffaele Sommese*, Gautam Akiwate‡, Mattijs Jonker*,
Giovane C. M. Moura§, Marco Davids§, Roland van Rijswijk-Deij*,
Geoffrey M. Voelker‡, Stefan Savage‡, K.C. Claffy†‡, Anna Sperotto*
* *University of Twente*, Enschede, the Netherlands
† *CAIDA* / ‡ *UC San Diego*, La Jolla, CA, USA
§ *SIDN Labs*, Arnhem, the Netherlands

***Abstract*—Anycast has proven to be an effective mechanism to enhance resilience in the DNS ecosystem and for scaling DNS nameserver capacity, both in authoritative and the recursive resolver infrastructure. Since its adoption for root servers, anycast has mitigated the impact of failures and DDoS attacks on the DNS ecosystem. In this work, we quantify the adoption of anycast to support authoritative domain name service for top-level and second-level domains (TLDs and SLDs). Comparing two comprehensive anycast census datasets in 2017 and 2021, with DNS measurements captured over the same period, reveals that anycast adoption is increasing, driven by a few large operators. While anycast offers compelling resilience advantage, it also shifts some resilience risk to other aspects of the infrastructure. We discuss these aspects, and how the pervasive use of anycast merits a re-evaluation of how to measure DNS resilience.**

***Index Terms*—DNS, Anycast**

I. Introduction

The architecture of the Domain Name System (DNS) is designed to distribute both load and responsibility. By delegating authority over distinct sub-trees of the DNS namespace, no single failure can disrupt the entire system. However, a nameservice failure at any particular node in the DNS hierarchy has the potential to disrupt access to all subordinate names (*i.e.,* if Verisign, the provider of nameservice for `.com`, were to fail, then queries to find the authoritative nameservers for `amazon.com`, `google.com` and others would also start to fail). DNS providers — and particularly those providing authoritative service for TLDs (top-level domains) or for large numbers of domains — must design their name server architecture to be resilient to failure.

Traditionally, this resilience relies on explicit name server replication. An authoritative nameserver provides a set of name server replicas in response to a query (*e.g.,* `ns1.foo.com`, `ns2.foo.com`, `ns3.foo.com`). If any such server fails — even silently — a requesting resolver can re-issue their request to a different replica. Distributing these name server replicas in disjoint networks insulates the overall service from the failure of any one network. So long as any one replica remains operational and reachable, name service can still be provided.

Over time another mechanism has emerged for providing resilience at the network layer: IP Anycast. In the IP anycast model, geographically diverse server replicas all use the *same* IP address by arranging for different networks to all announce the *same* network prefix. When a client sends a packet to the server's IP address, the packet will automatically be routed to the (topologically) closest replica. If the network connecting this replica fails, normal Internet routing processes will re-route packets to the next closest replica.[1] Employing this architecture for the DNS moves the decision for replica selection from an explicit choice made by the requesting party (typically a client or recursive resolver) to an implicit choice implemented by BGP and the ISP's routing policy.

In this paper, we focus on the evolution of anycast for providing DNS nameservice. We empirically characterize the adoption of anycast by nameservers supporting TLDs and SLDs (second level domains, *a.k.a.* registered domains), between 2017 and 2021. We show that anycast is now the dominant mode for providing DNS service — used by 97% of TLDs and 62% of SLDs in our dataset. We find that this adoption is not driven by the actions of individual domain owners, but is dominated by the engineering choices of a few large DNS infrastructure providers. The top 10 anycast-supporting DNS providers account for 92% of all domains with anycast nameservice. A single registrar, GoDaddy, accounts for the majority of anycast adoption in SLDs. To investigate the relationship between resilience and infrastructure diversity, we show that domains using anycast nameservice frequently exhibit lower diversity in their use of IP addresses and ASNs. As a result, anycast-based name service does not eliminate the resilience problem, but offers a different resilience risk profile. We conclude by reviewing different failure modes that shed light on how anycast changes the risk profile of a given deployment.

II. Related Work

The DNS ecosystem has been the subject of several studies focusing on diverse aspects of DNS resilience and robustness. Allman investigated the extent to which DNS administrators do not provide significant infrastructure diversity in hosting their domains [1]. Lame delegations also affect the robustness

[1]This failover assumes that the network failure leads to a route withdrawal. If the server fails, or the network fails silently (*e.g.,* such as from a DDoS attack), a routing update may not occur (Section VI).

978-3-903176-40-9

of the DNS ecosystem. Akiwate *et al.* found that lame delegations are surprisingly common, even in popular domains [2]. The consolidation of the Internet and DNS ecosystem, which we also observe in anycast adoption, has also been studied. Kashaf *et al.* found a considerable concentration in the use of third-party services for authoritative DNS nameservice [3]. Moura *et al.* found an increasing consolidation of the recursive resolution infrastructure [4]. The impact of operator practices on DNS query performance has also been extensively studied [5]–[8] culminating in recommendations for large DNS operators [9]. Our work focuses on anycast adoption in the authoritative nameserver infrastructure – an aspect that has not been investigated in depth. The use of anycast in DNS was first studied by Xun *et al.* [10] who used CHAOS queries to enumerate anycast instances, and estimate adoption of anycast in TLD authoritative nameservers. Their findings show, in 2013, between 56% to 72% of TLD authoritative nameservers adopted anycast. Our work expands their analysis by using anycast census data showing an increased adoption of anycast, in 2021, by 97% of TLDs. In 2015, Cicalese *et al.* [11] performed an anycast census using a methodology, called *iGreedy*, based on the Great-Circle Distance. Bian *et al.* [12] proposed a passive approach to anycast enumeration using public BGP data from route collectors. Recently, Sommese *et al.* proposed a methodology [13] to measure anycast using anycast vantage points. To make this study possible, our work uses measurements from Cicalese *et al.* and Sommese *et al.* (Section III). As such, our work leverages and builds on top of previous anycast enumeration studies.

III. Datasets

In this section we describe the DNS and anycast datasets that we use for our study and we outline considerations in handling this data.

A. Datasets

a) Anycast: Our work builds on two anycast IPv4 censuses. The first census was published in June 2017 by Cicalese *et al.*, who ran their *iGreedy* measurement on PlanetLab and RIPE Atlas to create an anycast census [11]. *iGreedy* uses the Great Circle Distance for anycast detection, enumeration and geolocation. More specifically, the technique uses speed of light violations to infer distinct anycast replicas and then uses multiple observations and a subsequent greedy algorithm to enumerate replicas. City-level geolocation relies on a maximum likelihood estimator. The resulting dataset contained 5486 distinct /24 anycast prefixes.[2]

Leveraging our previous work, we performed an anycast census in January 2021 using MAnycast^2 [13] to use in this study.[3] MAnycast^2 improves upon *iGreedy* with a filtering step that shortens measurement time considerably. MAnycast^2 relies on the principle of *using anycast to measure anycast*, which involves sending probes from multiple anycast vantage points to a target IP address and then checking which vantage points received the responses. The number of vantage points receiving responses reveals whether a target is unicast or anycast. MAnycast^2 uses *iGreedy* (on RIPE Atlas VPs) to cross-validate detected anycast prefixes and perform enumeration and geolocation. MAnycast^2 used 20 distinct vantage points provided by SIDN Labs. The January 2021 anycast census dataset contained 9999 distinct /24 anycast prefixes.

[2]iGreedy anycast census: https://anycast.telecom-paristech.fr/dataset/

[3]MAnycast^2 anycast census: https://github.com/ut-dacs/Anycast-Census/

Date	Total SLDs	Responsive SLDs	Unresponsive SLDs
2017-06-01	189.6 M	164.4 M (87%)	25.2 M (13%)
2021-01-31	210.4 M	187.5 M (89%)	22.9 M (11%)

TABLE I: Total SLDs and SLDs with responsive authoritative nameservers.

At the time of writing, there is no publicly available IPv6 anycast census. Hence, IPv6 anycast DNS infrastructure is out-of-scope for this work. As future work, we plan to expand our previous efforts [13] to also support IPv6 anycast measurement, and to compare IPv4 and IPv6 anycast authoritative nameserver deployments. Given that IPv6 is also popular among large companies, we expect that they often implement and offer anycast for both IP stacks.

b) DNS: We use DNS data provided by the OpenINTEL project, which measures ∼65% of the global DNS namespace by actively querying for the resource records of second-level domains (SLDs) under a sizable number of top-level domains (TLDs) on the Internet [14]. OpenINTEL's *daily* measurement actively queries for, among others, the authoritative nameserver records (i.e., NS records), as well as the IPv4 addresses (i.e., A records) of the names encountered in NS records. While the OpenINTEL project regularly expands its coverage of the namespace and has steadily added TLDs over time, we include only TLDs that were already covered at the time of the first anycast census in June 2017. This set consists of 1053 TLDs. To account for missing data points on particular days (which is rare but could occur, *e.g.,* due to incidental outages) we require each TLD to be in the dataset for 95% of all days between the anycast census dates. The TLDs we consider involve: the (legacy) generic TLDs `.com`, `.net` and `.org`; the new generic TLDs (ngTLD) such as `.tokyo`; and the country-code TLDs (ccTLD) `.at, .ca, .dk, .fi, .nl, .nu` and `.se`. The resulting DNS dataset accounts for ∼164 million domains in 2017 and ∼187 million in 2021.

To correlate a domain's anycast deployment with its popularity, we also use OpenINTEL measurement data for domain names of the top 1 million popularity lists for Alexa (2017-06-01 and 2021-01-31) and Cisco Umbrella (2021-01-31) .

c) Metadata: We use CAIDA's prefix-to-AS dataset [15] to map IP addresses of authoritative nameservers to their covering prefix and announcing AS number(s), and CAIDA AS-to-organization data [16] to map AS numbers to organizations. Finally, we use Netacuity data to geolocate unicast IPv4 addresses.

	2017-06-01	2021-01-31
TLDs	1533	1505
removed	–	60
added	–	32
∩	1473	
ccTLDs	247	
gTLDs	7	
ngTLDs	1219	

TABLE II: Root Zone TLD Snapshots in 2017 and 2021. Our work analyzes TLDs present in both 2017 and 2021 snapshots. Breakdown of the TLDs analyzed as either legacy gTLDs, new gTLDs or ccTLDs.

	gTLD		ccTLD		new gTLD		Total	
	2017	2021	2017	2021	2017	2021	2017	2021
Unicast	1	1	79	34	25	15	105(7.1%)	50(3.4%)
Mixed	2	1	137	160	117	139	256(17.4%)	300(20.4%)
Anycast	4	5	31	53	1077	1065	1112(75.5%)	1123(76.2%)
Total	7	7	247	247	1219	1219	1473	1473

TABLE III: Breakdown of TLDs with unicast, anycast, or mix of both anycast and unicast (mixed) authoritative nameservers in 2017 and 2021. Anycast adoption (including mixed) in 2021 reached $\sim 97\%$

B. Data considerations

Our analysis involves a few assumptions and decisions that factor into our results. First, we consider only responsive SLDs (Table I). Consequently, our results interact with active DNS infrastructure. We observed ∼12% unresponsive SLDs, which is consistent with the findings of Akiwate *et al.* [2].

Second, our analysis involves active DNS measurement data. Consequently, we consider nameservers learned from explicit NS queries, and A records that follow active resolution. We do not rely (directly) on the NS records and the A records (glue) in zone files. In other words, we consider only the records that are provided by the authoritative nameservers. Several studies have found inconsistencies between parent and child zones for up to 5–12% of observed SLDs [2], [17]. Whether DNS resolution follows parent or child records is resolver-dependent [17].

Finally, we are aware from the associated papers that both anycast inference methodologies we use can include classification errors [11], [13]. MAnycast^2 (combined with *iGreedy*) as well as *iGreedy* alone can result in false negatives (i.e., anycast deployments identified as unicast). However, both techniques deliver a conservative lower bound estimate of anycast deployment. Therefore, our anycast adoption analysis is a conservative lower bound estimation.

Our analysis relies on data collected on a daily (OpenINTEL) and quarterly basis (MAnycast^2). We developed the analysis code to be reused for reproducibility and continuous assessment of DNS anycast adoption. We publicly released the code for TLDs adoption analysis (§IV) [18]. Cases that sporadically require additional measurements (e.g., traceroute in §VI) are analyzed manually.

IV. Anycast adoption by TLDs

Given their critical role in the DNS, we start by characterizing anycast adoption by top-level domains (TLDs). We used snapshots of the root zone from DNS-OARC [19] for our two time periods. Table II summarizes the number and kinds of TLDs in each period. The total number of TLDs in the root zone decreased slightly from 1533 TLDs in 2017 to 1505 in 2021: 60 TLDs from 2017 were no longer in the 2021 root zone (such as `.intel`, `.telefonica`), and 32 new TLDs were added between 2017 and 2021 (*e.g.,* `.ss`, South Sudan's ccTLD, and `.amazon`). Between the two periods, though, most of the TLDs (1473) were *delegated* in both root zone files (∩).

We focus on this intersection of TLDs that are present in both zone files. Since there is significant variation in the types, history, management, and use of top-level domains [20], [21], we classify these TLDs into three categories: ccTLDs (*e.g.,* `.jp` and `.de`), gTLDs ("original" gTLDs: `.com`, `.edu`, `.gov`, `.mil`, `.org`, `.net`, `.int`) and ngTLDs (`.tokyo`, `.xyz`, `.top`).

For each TLD, we extract its NS records and associated A records. For each A record, we label it as using anycast if it matches the anycast prefix datasets described in §III-A, otherwise we label it as unicast. Since not all A records for a TLD may have the same label, we classify each TLD into three categories: those whose A records are all anycast, those whose A records are all unicast, and those with mixed usage (some, but not all, A records have anycast IPv4 addresses).

A. Increasing adoption of anycast

Table III shows the results of this TLD classification. Overall, the use of anycast for TLD authoritative nameservers, in whole or part, shows increased adoption between 2017 and 2021. In 2017, 1368 TLDs (93%) used anycast (in whole or in part), and just 105 (7%) used unicast. In 2021, anycast adoption increased, with 1423 TLDs (97%) using anycast while only 50 TLDs (3%) relied on unicast authoritative nameservers. For the ccTLDs, 45 of the 79 ccTLDs using unicast (57%) in 2017 moved to either mixed (23) or full (22) anycast infrastructure by 2021. For the ngTLDs, which already had widespread anycast adoption, 10 of the remaining 25 TLDs (40%) using unicast in 2017 moved to mixed or full anycast by 2021. For the original gTLDs, `.gov` moved from mixed to full anycast support, leaving `.mil` as the only original gTLD not using anycast.

A significant reason for the increase in ccTLDs using full anycast was the set of 18 ccTLDs in the mixed category in 2017, including `.cz`, `.io`, `.nl` and `.in`, that solely used anycast by 2021. For most of these (14 ccTLDs), the change was simply because they dropped the unicast nameserver. Perhaps 2017 marked a transition period where they balanced old and new infrastructure, and by 2021 those ccTLDs were committed to full anycast.

Not all changes increased anycast adoption. For instance, `.ki` (Kiribati) changed from full anycast to mixed infrastruc-

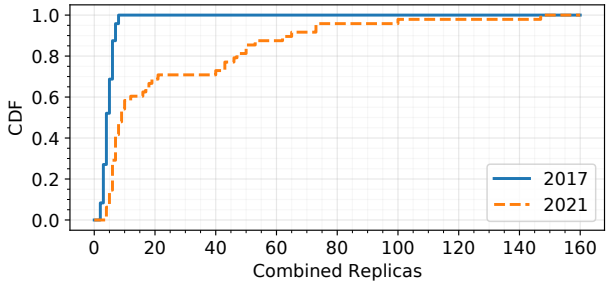

Fig. 1: Marked increase in combined replicas for ccTLDs moving to Mixed or Anycast from Unicast

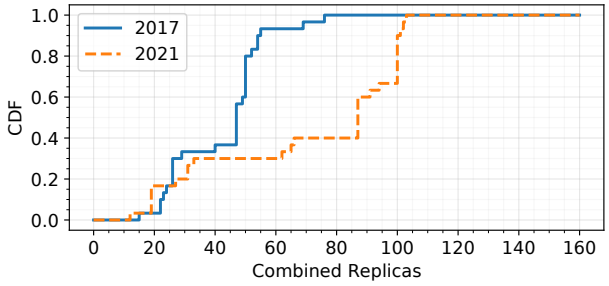

Fig. 2: Significant growth in replica infrastructure for ccTLDs using Anycast in both periods

ture, and three ccTLDs (`.ve`, `.pa`, and `.cd`)[4] changed from mixed to unicast only. These changes reflect the choice of, and dependence on, underlying services. The `.pa` and `.ve` ccTLDs, for example, employed the Internet System Consortium (ISC) authoritative anycast service (`sns-pb.isc.org`), which shut down on January 31, 2020 [22].

B. Anycast infrastructure expansion

One motivation for a TLD to switch from unicast to mixed or full anycast for authoritative name service is to improve availability and resilience. A proxy metric, admittedly rough, for representing the implications of this change is the scale of authoritative nameserver infrastructure. For a TLD, the total infrastructure is the combined unicast IPv4 addresses and anycast sites across all of the A records attached to the NS records for the TLD.[5] We refer to this count as the number of *combined replicas* providing name service for the TLD. For instance, France's `.fr` in January 2021 had 4 NS records with 4 IP addresses (1 unicast and 3 anycast), and the anycast addresses were distributed across 68 sites. In this case, we consider `.fr` to have 69 combined replicas.

The combined replica metric is a rough proxy metric because, from the client point of view, there is a difference between a unicast address, which is globally reachable, and an anycast site, which is reachable only to the portion of users mapped to that site by BGP (*i.e.,* anycast "fragments" the IP address space [5]). Still, it does reflect the infrastructure investment supporting name service and an upper bound on availability and resilience (discussed further in §VI).

We first focus on the 48 ccTLDs that were unicast only in 2017 and changed to mixed or full anycast by 2021. Figure 1 shows the CDFs of the number of combined replicas across these TLDs for both snapshots in time. In 2017 these ccTLDs had a median combined replica of 4 (in this case, 4 NS records each with 1 IPv4 address) and a 75%-ile of 6. After switching to anycast, the median number in 2021 increased to 9 replicas, and the 75%-ile to 43 — a significant increase from 2017 in terms of supporting infrastructure.

[4]Venezuela, Panama, and the Democratic Republic of Congo, respectively.

[5]We note that, due to the complexity of geolocating (and enumerating) anycast sites, we consider these results a lower-bound estimation.

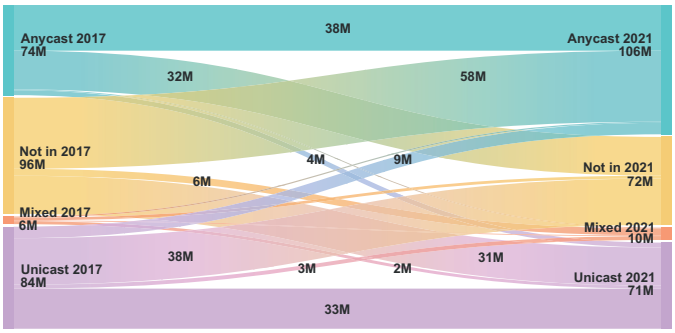

Fig. 3: Evolution of anycast authoritative deployment between 2017 and 2021.

For the new gTLDs, most (1025) used anycast in both 2017 and 2021, and only 10 switched from unicast to mixed or full anycast. We do not include a graph for these 10 new gTLDs, but to summarize they also experienced a significant increase in the scale of infrastructure: their median combined replicas increased from 4.5 to 34.

Moreover, Figure 2 shows that the scale of anycast infrastructure increased considerably between 2017 and 2021. The graph focuses on the 30 ccTLDs that used full anycast in both 2017 and 2020, and it shows the CDFs of the number of anycast sites across those TLDs. The median combined replicas increased from 47 in 2017 to 87 in 2021. Anycast infrastructure is scaling considerably over time, and name service naturally benefits from this scaling.

C. Large providers drive anycast adoption

For gTLDs and ngTLDs, the top 10 providers for TLDs account for 88% of the use of anycast by TLDs in 2021, with Neustar (36%), Afilias (22%) and Verisign (10%) leading the list. For ccTLDs, the landscape is more fragmented, reflecting the more nuanced balance that ccTLDs make between using first and third-party infrastructure. The top 10 providers are responsible for 69% of anycast ccTLDs, and the biggest operator, PCH, manages 25% of the ccTLDs, followed by NetNod (9%) and RIPE (5%).

Type	2017 Anycast	2017 Unicast	2017 Mixed	2017 Total	2021 Anycast	2021 Unicast	2021 Mixed	2021 Total
#SLD	74.2M (45.1%)	84M (51.1%)	6.2M (3.8%)	164.4M	106.4M (56.8%)	70.9M (37.8%)	10.2M (5.4%)	187.5M
#NS(IP)	10700	899028	N/A	909728	18179	756459	N/A	774638

TABLE IV: Anycast in DNS: 2017 and 2021 Adoption. Number of SLDs relying on unicast infrastructure decreased between 2017 and 2021 of 13.3%. Overall anycast adoption (Mixed+Anycast) reached 62.2%

V. Anycast adoption by SLDs

The TLD authoritative nameserver infrastructure has substantially adopted anycast. The next step in the resolution process is at the second-level domains (SLDs). Have SLDs followed suit, making DNS resolution fully reliant on anycast authoritative nameservers? To identify anycast authoritative infrastructure in SLDs, we correlate the 2017 and 2021 anycast census datasets with the authoritative infrastructure measurements provided by OpenINTEL. We extract the NS records of the SLDs and all of their related A records from OpenINTEL. Based upon the IPv4 addresses in their associated A records, for each domain we use the anycast census datasets to classify it as unicast, anycast, or a mix of the two.

Table IV summarizes the results of our classification for the 2017 and 2021 snapshots. As with TLDs, SLD infrastructure is also increasingly relying upon anycast infrastructure for authoritative name service. In 2017, 51.1% of domains relied on unicast infrastructure, 45.1% domains relied solely on anycast, and 3.8% of domains relied on a mix of the two. By 2021, the domains relying upon unicast dropped by 13.3%, the domains relying upon anycast increased by 11.7%, and the domains on mixed infrastructure increased by 1.6%.

Finally, looking at the IPv4 addresses of the authoritative servers, only 2.3% of them rely on anycast infrastructure. These results suggest that anycast is used by only a few companies, yet half of the domains in the DNS rely on these companies (§V-A).

To visualize the evolution of anycast adoption between 2017 and 2021, Figure 3 is a Sankey diagram[6] showing how domains changed categories between the two snapshots. For example, of the 84M domains relying on unicast in 2017, 33M of them still relied on unicast in 2021, 3.2M relied on a mix of unicast and anycast, 9.4M relied solely on anycast, etc. Given that anycast adoption is more prevalent in 2021, the diagram provides more detail on the sources of those domains. For instance, new domains are more likely to rely on anycast: nearly twice as many new domains in 2021 that did not exist in 2017 use anycast than unicast.

In contrast, the majority of SLDs that are no longer responsive in 2021 were using unicast infrastructure in 2017. Moreover, 9.4M domains shifted from a unicast infrastructure to full anycast, 3.2M from unicast to mixed anycast, and 4M shifting from anycast back to unicast. Examining these last 4M SLDs more closely, they are primarily domains with GoDaddy, Dyn, CloudFlare, and 1&1 moving to other minor registrars/infrastructures.

[6]Interactive Visual: https://public.flourish.studio/visualisation/5568561/

Org	SLD	%	Org	SLD	%
GoDaddy	44145357	54.53%	Google.	3433523	4.24%
CloudFlare	6955596	8.59%	Uniregistry	2376567	2.94%
1&1 IONOS	4808600	5.94%	Akamai	1451470	1.79%
DynDNS	3883403	4.80%	Amazon	1068653	1.32%
VeriSign	3878585	4.79%	One.com	1016796	1.26%

TABLE V: Top 10 Anycast Organizations 2017, responsible for 90% of the anycast adoption. GoDaddy was market leader.

Org	SLD	%	Org	SLD	%
GoDaddy	52681291	44.11%	1&1 IONOS	6033089	5.05%
Cloudflare	15252317	12.77%	NSONE	3160888	2.65%
Google	11014408	9.22%	Amazon	2949373	2.47%
NeuStar	7968959	6.67%	NetActuate	1902258	1.59%
Zenlayer	6800764	5.69%	Tencent	1781520	1.49%

TABLE VI: Top 10 Anycast Organizations 2021, responsible for 92% of the anycast adoption. GoDaddy's market share slightly decrease, Cloudflare increased.

A. Concentrated set of providers drives anycast adoption

The increasing adoption of anycast among TLDs is tied to the deployment of anycast by a concentrated set of providers (§IV-C), and we similarly look at providers to explain increased anycast adoption among SLDs. Using IP→AS→organization mappings (§III-A), we identified the top 10 anycast organizations both in 2017 (Table V) and in 2021 (Table VI). These top 10 anycast organizations are responsible for 90% of anycast adoption in 2017 and ∼92% in 2021. These results confirm that adoption is primarily driven by large DNS providers.

Looking at individual companies, GoDaddy unsurprisingly is the largest company by far, in terms of SLDs hosted, that operates anycast services for their authoritative nameservers. In 2017, GoDaddy accounted for more than half (∼55%) of the anycast SLDs. In 2021 the percentage decreased (∼44%), but the absolute numbers increased. GoDaddy is the largest registrar in the world, and therefore their infrastructure choices as registrar (and, by default, DNS hosting provider) heavily influence the DNS ecosystem.

Next is Cloudflare, where anycast adoption for SLDs increased from ∼9% in 2017 to ∼13% in 2021. In contrast to GoDaddy, Cloudflare's core business is not as a registrar (even if it recently started a registrar service), but to offer CDN and DDoS protection services to their customers. As a result,

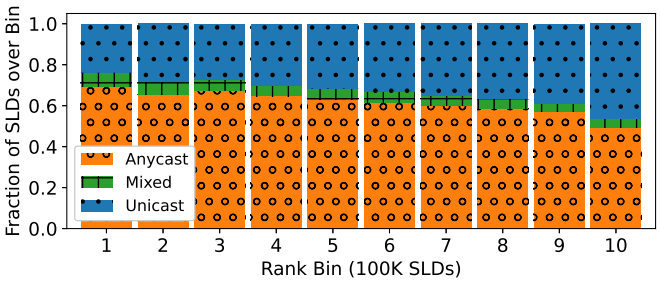

Fig. 4: Anycast adoption correlated with domain popularity as ranked by Cisco Umbrella.

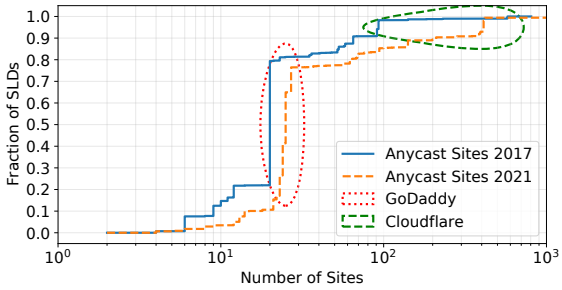

Fig. 5: Distribution of number of anycast authoritative nameserver sites per SLD. In 2021, average number of sites slightly increased.

customers likely choose Cloudflare for better performance, resilience, and availability of their Web services. But since Cloudflare adopted anycast, its customers benefit from it as well. In short, technical and business decisions of the company drive anycast adoption.

Among the other top 10 providers is a mixture of popular Web site building and hosting (1&1) and cloud providers, which operate DNS hosting themselves (*e.g.,* Route53 and Cloud DNS) or with third parties (*e.g.,* other DNS registrars).

In contrast to providers using anycast, unicast deployment is less concentrated: the top 10 account for only 63% of the total unicast SLDs. In terms of types of companies, the top 10 unicast DNS providers, both in 2017 and 2021, are almost all Chinese providers with two notable exceptions of Amazon and OVH. For Amazon, nearly 6.6 million SLDs were hosted on non-anycast services (primarily third-party EC2 instances). OVH, a popular European hosting provider, offers optional anycast service for DNS nameservers, and customers must pay a premium of €1.21/year for anycast. Nearly all SLDs using OVH's authoritative infrastructure use unicast: we measured 4,156,201 domains using OVH's unicast infrastructure, and just 130,951 domains using its anycast infrastructure. We speculate that offering anycast as an optional paid service results in low anycast adoption for OVH customers.

B. Role of registrars in anycast adoption

The GoDaddy example shows that popular registrars play a fundamental role in anycast adoption. Popular registrars like GoDaddy generally operate across the entire gTLD market, resulting in a roughly similar degree of anycast adoption from the new gTLDs (55.4%) to `.com`, `.net`, and `.org` (64%).

The ccTLD perspective looks quite different, with a much lower overall adoption of 37.3% of SLDs. One example particularly stands out, where the anycast adoption in `.se` (Sweden) is notably high while the anycast adoption in `.nl` (Netherlands) is comparatively low. The adoption of anycast in `.se` is related primarily to the implementation by Loopia AB, the largest registrar in Sweden (confirmed by IIS [23]). The largest registrar in the Netherlands, TransIP B.V., has yet to adopt anycast. More generally, `.nl` domains are spread across different small local registrars, which usually do not want to implement a global anycast infrastructure (due to related costs and complexity) or to pay for third-party service. These factors accentuate the low adoption of anycast for `.nl` domains.

In short, as expected large registrars play a fundamental role in anycast adoption, with GoDaddy for gTLDs and Loopia for ccTLDs serving as notable examples.

C. Anycast adoption and domain popularity

Popular domains by necessity are scalable, reliable, and available, and anycast is an increasingly popular mechanism to support those goals. As a result, either due to the extensive infrastructure that the domains deploy themselves, or by relying upon large-scale third-party infrastructure, we expect domain popularity to correlate with the use of anycast for the domain's authoritative name service.

To validate this expectation, we check which SLDs on the Cisco Umbrella list on January 31, 2021, relied upon anycast infrastructure for their nameservers.[7] In particular, we group sets of 100,000 ranked SLDs together into bins. We then calculate the percentage of SLDs in each bin that rely upon anycast using the 2021 anycast census dataset. For each ranked bin, Figure 4 shows the fraction of SLDs that use unicast, anycast, or a mix of the two for authoritative name service. The results clearly confirm the expected correlation that more popular domains rely upon anycast. Indeed, among the 100,000 highest ranked domains, more than 76% of the SLDs rely upon anycast in whole or part.

Since the use of anycast for authoritative name service often depends on the underlying provider, we also examine the providers behind the most popular domains. Looking at the top 10,000 SLDs of the Umbrella list, there are many different anycast providers (*e.g.,* Google, Facebook, Microsoft, Apple, etc.) together with classic CDN and DDoS protection providers such as Cloudflare, Amazon, Akamai, etc. Looking at the entire top 1M Umbrella list, Cloudflare (and partially Amazon) lead the anycast adoption market. GoDaddy accounts for 10% of the SLDs using anycast, and those SLDs tend to be less popular. While the extensive DNS scan showed GoDaddy to be the most popular anycast provider due to its DNS market

[7]We used Umbrella instead of Alexa since Umbrella ranks domains based on the DNS query load received by the Cisco Umbrella OpenDNS service. As a result, this measure correlates better with popularity in DNS resolution.

Source	2017 Anycast	2017 Mixed	2017 Unicast	2017 Total	2021 Anycast	2021 Mixed	2021 Unicast	2021 Total
.com	53.3M (49.2%)	3.9M (3.6%)	51.2M (47.2%)	108.3M	76.7M (58.7%)	8.2M (6.3%)	45.7M (35.0%)	130.7M
.net	5.7M (45.1%)	0.5M (3.7%)	6.5M (51.2%)	12.6M	6.3M (54.3%)	0.6M (4.9%)	4.7M (40.8%)	11.6M
.org	4.9M (53.4%)	0.3M (3.2%)	4.0M (43.4%)	9.3M	6.0M (64.0%)	0.2M (1.6%)	3.2M (34.4%)	9.4M
ngTLDs	7.5M (34.7%)	1.2M (5.4%)	13.0M (60.0%)	21.7M	12.4M (55.4%)	1.0M (4.5%)	9.0M (40.1%)	22.4M
ccTLDs	2.6M (21.8%)	0.5M (3.7%)	9.0M (74.5%)	12.1M	4.9M (37.3%)	0.2M (1.5%)	8.0M (61.2%)	13.1M
.se	614K (43.3%)	146K (10.3%)	660K (46.5%)	1421K	810K (57.2%)	13K (0.9%)	594K (41.9%)	1416K
.nl	256K (4.8%)	70K (1.3%)	5014K (93.9%)	5400K	1277K (22.2%)	36K (0.6%)	4446K (77.2%)	5760K
Alexa	337K (35.3%)	33K (3.4%)	584K (61.3%)	953K	423K (51.6%)	14K (1.7%)	383K (46.7%)	820K
Umbrella	N/A	N/A	N/A	N/A	157K (61.1%)	13K (4.9%)	87K (33.9%)	256K

TABLE VII: Anycast in DNS: 2017-2021 Adoption per TLD. The specific case of the Netherlands and Sweden shows how two similar countries can have a completely different anycast adoption for authoritative nameservers due to registrar choices.

share (§V-A), the bulk of its customer base has SLDs that are in the long tail of the popularity distribution.

D. Anycast infrastructure expanding

The anycast datasets include the number of anycast sites for each anycast IPv4 address. As in §IV-B for TLDs, we can use this information to examine how the scale of the underlying anycast infrastructure has changed over time for SLDs.

Figure 5 shows CDFs of the number of anycast sites supporting authoritative name services across all SLDs. The graph focuses on just the SLDs fully using anycast, and for each SLD we sum the number of anycast sites across all IPs associated with the SLD. Since GoDaddy is the most common provider of anycast for SLDs (§V-A), it determines the largest mode of the distribution. In 2017, GoDaddy used 20 distinct anycast sites, making this scale the most common anycast deployment for a domain. By 2021 GoDaddy had expanded its infrastructure slightly to 25–30 anycast sites.

In contrast, Cloudflare significantly expanded its anycast infrastructure between 2017 and 2021. In 2017 SLDs relying upon Cloudflare were supported by ∼90 anycast sites, and by 2021 the number of sites increased to 130.

VI. Implications of Anycast Adoption for DNS Resilience Risk Profiles

The goal of the redundancy mechanism in the DNS is resilience against failure. In traditional unicast DNS, the recommended best practice is to maintain at least two authoritative nameservers (IP diversity) in different network segments (routed prefix diversity), and ideally in different networks (AS diversity) and geographic regions (geographic diversity) [24]. This investment in diversity provides resilience against failures of individual servers, subnets, entire networks, or connectivity in a specific region.

With unicast, the settings explicitly manifest this diversity. If one server fails, the client resolver can (and must) be responsible for re-issuing the query to a different authoritative nameserver. However, it can be operationally costly and complex to arrange for the subnet diversity recommended for unicast DNS deployments. We speculate that the complexity of arranging topological and geographical diversity for authoritative DNS is a major driving force behind the introduction of large anycast services where domain registrants can outsource their authoritative DNS service provisioning.

With anycast, service diversity is not explicit in the DNS settings, but manifests in the routing system. That is, in case of a link failure of a single authoritative nameserver, the Internet routing processes re-route packets to a different authoritative nameserver replica with the same IP address. Importantly, the effect of anycast adoption on resilience depends on deployment parameters as well as failure conditions. For example, if a server fails, the client can go to another server. If the IPs are diverse, then a failure of one network can be tolerated by a client-side retry. If the IPs are routed by different ASes then the same mechanism tolerates ISP-level failures.

Anycast hides at least some of the replica choice decision from the client. If all NS entries point to the same IP, then resolution relies entirely on anycast and if a server or subnet fails silently (*i.e.,* there is no route withdrawal) then everyone routed to that advertiser is effectively black-holed. If the domain is both using anycast and returns multiple replicas, then the client can still tolerate failure via retry so long as those distinct replicas are not all a) the same IP (in case of server-level failure), b) on the same network (network-level failure), or on the same AS (AS-level failure). Potentially, anycast re-routing could interrupt TCP sessions, which are usually used by DNS resolvers when responses exceed 512 bytes and EDNS is not supported. However, we advocate that the risk and the possible impact is small, given the short-lived TCP sessions and the in-protocol reliability mechanisms.

Notably, most SLDs rely on two authoritative server IPs both for anycast and unicast deployments (Figure 6), likely because registries and registrars usually require two distinct NS records. However, the number of distinct IPs for anycast deployments peaks at 8, and 90% have 4 or fewer. In contrast, 10% of unicast deployments have 12 distinct IP addresses. The AS-level diversity has a similar contrast: the vast majority of anycasted SLDs are anycasted entirely from a single AS or sibling ASes that belong to the same company (*e.g.,* Neustar), while 40% of unicast SLDs use two or more ASNs (Figure 7). Concentration of services within a single AS/company is a natural market force, but comes with its own risks. Manifestations of these risks [25] has motivated the

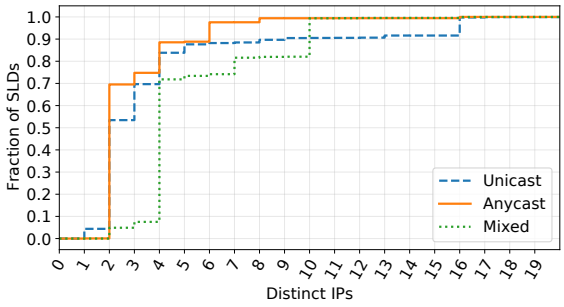

Fig. 6: Number of IPs corresponding to SLD authoritative nameservers. Anycast authoritative nameserver deployments tend to have fewer IPs, since anycast provides diversity via the routing system.

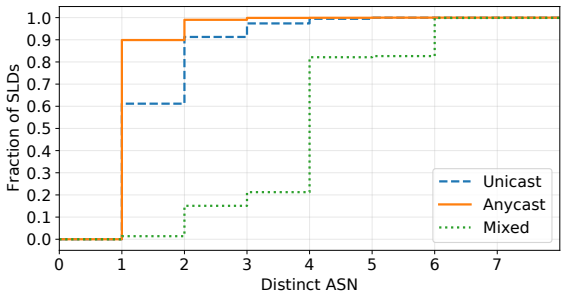

Fig. 7: AS Diversity for SLDs with anycast, unicast and mixed authoritative nameservers. Anycast deployments are usually concentrated in a single ASN

non-significant number of mixed deployments, that attempt to optimize multiple dimensions of DNS resilience [26].

We also found two interesting anycast cases of this "single point of failure", with many domains routed behind the same prefix: 1&1 IONOS SE, responsible for ∼6.8 million domains; and Loopia AB. 1&1 IONOS SE announces their anycast network from a single /22 block; our anycast geolocation data (based on the iGreedy second-stage measurement of [13]) indicates that all their anycast sites are in the same location for all the different IPs. Using traceroute from all the nodes of the RING NLNOG network troubleshoot platform [27], we discovered that the four /24 networks composing the 1&1 IONOS SE /22 block are routed behind the same last hop. Similarly, Loopia AB serves ∼500K domains via anycast from a single /24 block. This means that Loopia relies as its only resilience mechanism uniquely on anycast, with all the consequences related to the possible silent failing of one of the instances (*e.g.,* DNS unreachability for part of the users). The 5% of nameservers pointing to the same IP (Figure 6) is another interesting case of a single point of failure. Even if registrars require two nameservers, operators effectively provide lower diversity by pointing the nameservers to the same IP. We find this behavior also in ∼8% and ∼3% Alexa and Umbrella domains, shifted towards the tail of the lists.

Placing servers in different geographical locations reduces

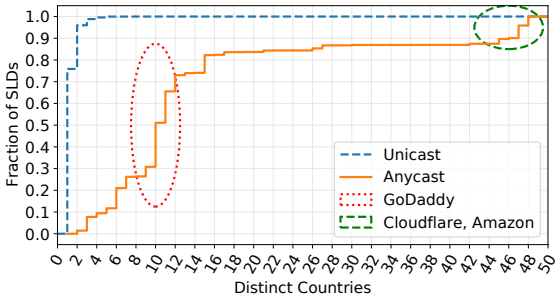

Fig. 8: Distribution of SLD authoritative nameservers across countries. Anycast deployments are more globally distributed.

latency and improves resilience against disasters. Given that anycast deployments are often globally distributed, we expect and observe higher country diversity (Figure 8) for these deployments. For unicast infrastructure, ∼75% of domains rely on authoritative servers in the same country, ∼20% spread over two countries and only ∼5% spread over more than two. In contrast, anycast authoritative services are hosted in 6–12 countries on average (depending on whether GoDaddy is the provider). Another ∼14% are hosted in more than 42 distinct countries (CloudFlare and Amazon deployments). To conclude, anycast can suffer from administrative or business failures (*e.g.,* global misconfiguration, attacks, etc) for centralized deployments (*i.e.,* single companies), but, at the same time, it helps to increase geographical availability and resilience of the DNS ecosystem.

VII. Conclusions and Future Work

We have characterized anycast adoption in authoritative DNS infrastructure for TLDs and SLDs. We found high adoption of anycast as a resilience mechanism, reaching 97% for TLDs and 62% for SLDs. This adoption is driven mostly by engineering choices of few very large DNS infrastructure providers. In our data set, one provider (GoDaddy) was responsible for the majority of anycast adoption in SLDs. Finally, we examined the relationship of anycast deployments to other traditional metrics of infrastructure diversity. Our findings show that anycast adoption changes the DNS service availability risk profile but does not eliminate all such risks. In fact, anycast can hide certain types of availability failures, and limit recovery options. A mixed deployment that includes traditional unicast redundancy as well as anycast options mitigates this risk, but increases cost and complexity.

As future work, we will focus on characterizing the resilience of anycast, unicast, and mixed deployments, including implications on performance metrics such as resolution latency. Moreover, we will study how different providers operate and optimize their anycast network. We will also examine DNS resolution behavior from a client perspective using different vantage points, showing how different failure modes (DDoS attacks, physical failure, misconfiguration) affect resilience of different deployment strategies.

Acknowledgments

We thank our shepherd Matteo Varvello and the TMA anonymous reviewers for their insightful suggestions and feedback. We also thank Alberto Dainotti and Ulrich Wisser for their valuable time, insights, and feedback. This work was supported in part by: the NWO-DHS MADDVIPR project (628.001.031/FA8750-19-2-0004); National Science Foundation grants CNS-1764055, CNS-1903612, OAC-1724853, CNS-1901517, CNS-1705050, and CNS-1629973; DARPA Coop. Agg. HR00112020014; and the EU H2020 CONCORDIA project (830927).

References

[1] M. Allman, "Comments on DNS Robustness," in *Proceedings of the Internet Measurement Conference*, 2018.

[2] G. Akiwate, M. Jonker, R. Sommese, I. Foster, G. M. Voelker, S. Savage, and K. Claffy, "Unresolved Issues: Prevalence, Persistence, and Perils of Lame Delegations," in *Proceedings of the ACM Internet Measurement Conference*, 2020.

[3] A. Kashaf, V. Sekar, and Y. Agarwal, "Analyzing Third Party Service Dependencies in Modern Web Services: Have We Learned from the Mirai-Dyn Incident?" in *Proceedings of the ACM Internet Measurement Conference*, 2020.

[4] G. C. M. Moura, S. Castro, W. Hardaker, M. Wullink, and C. Hesselman, "Clouding up the internet: How centralized is DNS traffic becoming?" in *Proceedings of the ACM Internet Measurement Conference*, 2020.

[5] R. de Oliveira Schmidt, J. Heidemann, and J. H. Kuipers, "Anycast latency: How many sites are enough?" in *Passive and Active Measurement*, 2017.

[6] G. Moura, J. Heidemann, W. Hardaker, J. Bulten, J. Ceron, and C. Hesselman, "Old but gold: Prospecting TCP to engineer DNS anycast (extended)," *Tech. Rep. ISI-TR-740*, 2020. [Online]. Available: https://www.isi.edu/%7ejohnh/PAPERS/Moura20a.html

[7] M. Müller, G. C. M. Moura, R. de O. Schmidt, and J. Heidemann, "Recursives in the wild: Engineering authoritative DNS servers," in *Proceedings of the 2017 Internet Measurement Conference*, 2017.

[8] K. Schomp and R. Al-Dalky, "Partitioning the internet using anycast catchments," *SIGCOMM Comput. Commun. Rev.*, vol. 50, no. 4, pp. 3–9, Oct. 2020.

[9] G. Moura, W. Hardaker, and J. Heidemann, "Considerations for Large Authoritative DNS Servers Operators," Internet Draft, Feb. 2021. [Online]. Available: https://tools.ietf.org/pdf/draft-moura-dnsop-authoritative-recommendations-08.pdf

[10] X. Fan, J. Heidemann, and R. Govindan, "Evaluating anycast in the domain name system," in *2013 Proceedings IEEE INFOCOM*, 2013.

[11] D. Cicalese and D. Rossi, "A Longitudinal Study of IP Anycast," *SIGCOMM Comput. Commun. Rev.*, vol. 48, no. 1, pp. 10–18, Apr. 2018.

[12] R. Bian, S. Hao, H. Wang, A. Dhamdere, A. Dainotti, and C. Cotton, "Towards passive analysis of anycast in global routing: Unintended impact of remote peering," *SIGCOMM Comput. Commun. Rev.*, vol. 49, no. 3, pp. 18–25, Nov. 2019.

[13] R. Sommese, L. Bertholdo, G. Akiwate, M. Jonker, R. van Rijswijk-Deij, A. Dainotti, K. Claffy, and A. Sperotto, "MAnycast2: Using Anycast to Measure Anycast," in *Proceedings of the ACM Internet Measurement Conference*, 2020.

[14] R. van Rijswijk, M. Jonker, A. Sperotto, and A. Pras, "A High-Performance, Scalable Infrastructure for Large-Scale Active DNS Measurements," *IEEE journal on selected areas in communications*, vol. 34, no. 6, pp. 1877–1888, Jun. 2016.

[15] CAIDA, "Routeviews Prefix to AS mappings Dataset for IPv4 and IPv6," 2020. [Online]. Available: http://www.caida.org/data/routing/routeviews-prefix2as.xml

[16] CAIDA, "Inferred AS to Organization Mapping Dataset," https://www.caida.org/data/as-organizations/, 2017.

[17] R. Sommese, G. C. Moura, M. Jonker, R. van Rijswijk-Deij, A. Dainotti, K. C. Claffy, and A. Sperotto, "When parents and children disagree: Diving into DNS delegation inconsistency," in *International Conference on Passive and Active Network Measurement*, 2020.

[18] "Root Analysis–Source Code," https://github.com/gmmoura/tma2021.

[19] DNS OARC, "Introduction to DNS-OARC," https://www.dns-oarc.net.

[20] T. Halvorson, M. F. Der, I. Foster, S. Savage, L. K. Saul, and G. M. Voelker, "From .Academy to .Zone: An Analysis of the New TLD Land Rush," in *Proceedings of the Internet Measurement Conference*, 2015.

[21] M. Korczynski, M. Wullink, S. Tajalizadehkhoob, G. C. M. Moura, A. Noroozian, D. Bagley, and C. Hesselman, "Cybercrime After the Sunrise: A Statistical Analysis of DNS Abuse in New GTLDs," in *Proceedings of the Asia Conference on Computer and Communications Security*, 2018.

[22] ISC, "Secondary Name Services," https://www.isc.org/sns-pb/, 01 2020.

[23] Swedish Internet Foundation, "Registrars .se," https://internetstiftelsen.se/en/domains/domain-statistics/registrars-se/, Apr. 2020.

[24] R. Elz, R. Bush, S. Bradner, and M. Patton, "Selection and Operation of Secondary DNS Servers," IETF, RFC 2182, Jul. 1997. [Online]. Available: http://tools.ietf.org/rfc/rfc2182.txt

[25] Scott Hilton, "Dyn Analysis Summary Of Friday October 21 Attack," https://web.archive.org/web/20190225060705/https://dyn.com/blog/dyn-analysis-summary-of-friday-october-21-attack/, 10 2016.

[26] A. Abhishta, R. van Rijswijk-Deij, and L. J. M. Nieuwenhuis, "Measuring the Impact of a Successful DDoS Attack on the Customer Behaviour of Managed DNS Service Providers," *SIGCOMM Comput. Commun. Rev.*, vol. 48, no. 5, pp. 70–76, Jan. 2019.

[27] Job Snijders, "NLNOG RING," https://ring.nlnog.net/.