Measuring the Practical Effect of DNS Root Server Instances: A China-Wide Case Study

Fenglu Zhang[1], Chaoyi Lu[1], Baojun Liu[1,4]✉, Haixin Duan[1,2,3] and Ying Liu[1]✉

[1] Tsinghua University [2] Qi An Xin Group [3] Peng Cheng Laboratory
[4] Beijing National Research Center for Information Science and Technology (BNRist)
{zfl20,lcy17}@mails.tsinghua.edu.cn
{lbj,duanhx}@tsinghua.edu.cn, liuying@cernet.edu.cn

Abstract. DNS root servers are deployed using multiple globally distributed anycast instances, and the scale of instances across the globe has been rapidly growing. This paper presents a measurement study that investigates the practical effect of root server instances deployed in the Chinese mainland. Our analysis of this issue includes two-fold. First, we measure the catchment area of the root server instances and answer the question about which domestic networks are served. Our results show that some of the instances are not accessible from major ISP networks due to limits of BGP routing policies, and a number of root queries still turn to further instances outside the international gateway. Second, we evaluate the impact of deploying new instances on query performance and root server selection in resolvers. We confirm that root instances contribute to lowered query delay from networks within their catchment area. Through reviewing source code of mainstream DNS implementations, we find that less-latent root servers are generally preferred thus deploying root server instances increase their possibilities to absorb DNS root requests from nearby resolvers. We make recommendations to improve the operational status of the DNS root server system.

1 Introduction

The DNS root is the starting point of the domain name space that bootstraps all DNS queries. To resist denial-of-service (DoS) attacks and improve stability of the DNS, all 13 root servers are deployed using *anycast* [18] with multiple *root server instances* that act collectively behind. Through BGP configuration, each DNS root query is routed to one instance, preferably the closest to its origin [34]. Network operators may host an instance, or peer with networks of root server operators to improve access to the DNS Root Server System (RSS) [13].

In recent years, there has been rapid deployment of new instances in the RSS, with an aim of providing faster and more reliable service to "*underserved*" areas (e.g., areas perceiving poor root query performance) [43]. At the time of writing, 1,469 root server instances are operating across the globe, which is 44.7% more compared to the number in 2019 [6]. However, information about their hosting networks and peering ASes is kept private, and the *practical effect* of the instances is still opaque to network operators. A set of research questions that are critical to understanding the deployment process are still not answered, including: *Which networks are within their catchment area? Which networks are still not served by close instances? How are they actually absorbing DNS root queries from nearby resolvers?* A few existing studies and tools focus on measuring delay to root servers [32, 23, 28], root manipulation [27] and the health status of root servers [5, 22, 42], but insights into the practical effect of root server instances is still lacking from the perspective of recursive resolvers.

We believe that seeking answers to the above questions helps examine the operational status of root server instances and can provide guidelines to their future deployment. In this paper, as a first step forward, we perform a case study on the practical effect of the 16 root server instances that have been deployed in the Chinese mainland[1], an understudied region with a large Internet population. Taking advantage of a side channel embedded in the DNS censorship mechanism [35], we propose a novel methodology to measure the catchment area of domestic root server instances (Section 3). While the instances do serve nation-wide areas, our results also show that some of them are not accessible from major ISP networks due to limits of BGP routing policies (Section 4). One `I-Root` instance even cannot be accessed from all three major ISP networks. As a result, a number of root queries still turn to further instances outside the international gateway, despite that 16 closer domestic instances are operating. Further, we investigate how deploying domestic instances can actually impact root server selection and absorb queries from recursive resolvers within their catchment area (Section 5). For this task, we measure the delay to root servers and review the source code of 4 types of common recursive resolver implementations: BIND 9 [25], Unbound [40], Knot Resolver [16] and PowerDNS Recursor [41]. We confirm that deploying domestic instances effectively lowers query delay, and their corresponding root servers will thus be preferred by the selection algorithms (especially in BIND and Knot Resolver).

From our findings, we make recommendations to multiple parties (including network operators and DNS software vendors) to improve the efficacy of the RSS. We believe that this work provides valuable insights into understanding the operational status of root server instances.

2 Background and Related Work

DNS Root Server System (RSS). Due to early payload size limits, there are only 13 root servers in the RSS [13], which are named by `A-Root` through `M-Root`. The 13 root servers are administered by twelve root server operators, such as Verisign and ICANN. All root servers in the RSS serve one unique copy of the DNS root zone managed by IANA [19]. To resist denial-of-service (DoS) attacks and improve stability of the DNS, all root servers are currently deployed using *anycast* [18] that allows multiple *root server instances* to act collectively

[1] Due to different network policies, in this paper we exclude Hong Kong SAR, Macao SAR and Taiwan from the scope of our study.

Table 1. Root server instances deployed in the Chinese mainland

Root	Global	Local	Geo-Locations
F-Root	0	4	Hangzhou, Beijing, Chongqing, Xining
I-Root	1	0	Beijing
J-Root	2	0	Beijing, Hangzhou
K-Root	1	2	Guangzhou, Guiyang, Beijing
L-Root	6	0	Beijing (2), Shanghai, Zhengzhou, Wuhan, Xining

using the same address. Over 1,400 instances are now operating in the RSS [6] and at the time of our experiment in this paper (Dec 2020), 16 instances have been deployed in the Chinese mainland. Table 1 shows their details.

Catchment area of root server instances. To improve their access to the RSS, local networks may *peer* to root server operator networks through exchanging BGP routing information [24, 21], or may apply to *host* a root server instance [26, 39, 20] in their networks or Internet exchange points (IXes). According to their different catchment area, the RSS comprises both *Global* and *Local* root server instances. Local instances only serve a limited network range and their catchment area is limited to the hosting ASes or the boundaries of BGP confederation [7, 34]. By contrast, Global instances let BGP alone determine their service scope. As listed in Table 1, the Chinese mainland hosts 6 Local and 10 Global instances.

Unauthorized root servers. Unauthorized root servers are those established outside of the RSS. As one type of DNS manipulation, operators of unauthorized root servers take control of the entire DNS name space in their service area. Previous studies have discovered one server potentially masquerading F-Root nodes in 2013 [17] and confirmed an unauthorized root mirror that exclusively serves CERNET (China Education and Research Network) in 2016 [27].

Related work. To understand the performance and security of the RSS, efforts have been devoted to investigating the impact of uneven distribution of root server instances on end-user query latency [32, 28], evaluating effects of anycast through examining DNS traffic and BGP data [12, 37, 34, 49], and detecting DNS root manipulation in the wild [17, 27]. However, little has been done to understand the practical effect of anycast instances behind root servers, or how their deployment and operation can be improved in the future.

In addition, a series of works examine how common DNS resolvers select and query authoritative servers (NSes, instead of root servers) [50, 38, 8]. Almost all of them answer this question by designing simulation experiments or inspecting outgoing DNS queries. They conclude most implementations prefer authoritative servers with the lowest latency, while others choose randomly. However, the reasons behind remain unrevealed, as few of them provide source code analysis.

3 Vantage Points and Methodology

In this section, we elaborate on how we collect vantage points that have broad coverage in the Chinese mainland, as well as our approach to measuring the catchment area of domestic root instances and delay to root servers.

3.1 Vantage point (VP) selection and validation

There are three major ISPs in the Chinese mainland, including China Telecom, China Unicom and China Mobile. Typically, the ISP networks are managed at a provincial level (there are 31 provinces in the Chinese mainland, excluding Hong Kong SAR, Macao SAR and Taiwan), and the network policies may differ in each province. However, common global measurement platforms (e.g., RIPE Atlas [44]) do not have good provincial coverage of VPs in China.

For our study, we select a Chinese commercial network looking glass platform that supports DNS queries. The platform operates over 300 VPs in all 3 major ISPs and multiple provinces, as well as in CERNET (China Education and Research Network) that serves universities. Each VP allows us to issue basic IPv4 DNS queries to custom server addresses, but does not offer additional DNS functionalities (e.g., NSID [10] that requests the identity of DNS server instances).

Since the advertised VP locations on commercial platforms cannot be relied on [48], we validate the locations of each VP before our experiment. After establishing a custom DNS server, we launch DNS queries from each VP to the server and check the source addresses of incoming queries against the MaxMind database [36]. If the locations do not match what they advertise, we remove the VPs from consideration. Meanwhile, to avoid DNS hijacking by middleboxes (e.g., NXDOMAIN rewriting [46]), on each VP we send DNS queries to 5 IP addresses that *do not* provide DNS service (i.e., normally, the queries will time out). We remove all VPs that receive DNS responses in the test and put detailed analysis on the filtered VPs in Appendix B. In the end, we select **182 vantage points** that advertise the correct location and are not affected by DNS hijacking. As shown in Table 2, they cover 31 provinces in the Chinese mainland. Due to the consideration of limited vantage points, we have to exclude the regional (province) differences analysis from the scope of our study.

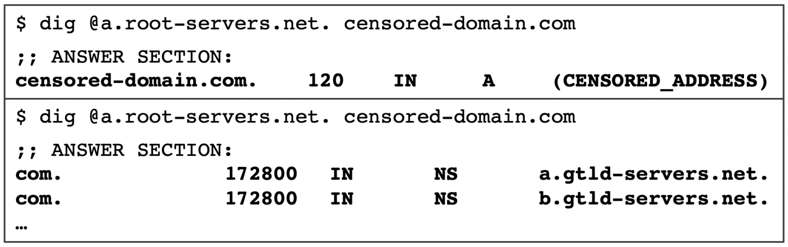
Table 2. Count and coverage of selected vantage points

ISP	# VPs	Provincial Coverage
China Telecom	71	26/31
China Unicom	74	28/31
China Mobile	24	21/31
CERNET	13	8/31
Total	**182**	**31/31**

3.2 Methodology

Catchment area of domestic root instances. The Chinese mainland hosts 16 root server instances (as listed in Table 1) and we seek to measure whether they are able to serve domestic networks. To find the exact instance that responds to a DNS server, one may issue NSID [10] or CHAOS-class queries [14] that return instance-specific strings (e.g., "`s1.ash`" represents the Ashburn instance of `I-Root`). A DNS traceroute [47] also reveals the path of root queries and gives clues about the destination instances. However, these DNS functionalities are not supported by most measurement platforms that offer broad ranges of Chinese VPs, including ours.

We try to overcome the challenges by posing another question: *are DNS queries from VPs in the Chinese mainland to root servers resolved domestically or overseas?* If resolved domestically, then the catchment area of domestic instances covers the networks of corresponding VPs. Fortunately, we find that this task is possible by leveraging the DNS censorship mechanisms of China [35], which are deployed at the international gateway to block access to certain websites [15] (we also confirm the location of censorship systems through offline discussions with large ISPs). On detection of DNS queries carrying censored domains (e.g., `google.com`), forged responses are injected before the authentic ones arrive [9]. As shown in Figure 1 (top), if a root query carrying a censored domain leaves the international gateway, its response will contain an `A` record pointing to blocked IP addresses [31]. Note that we have removed VPs that witness DNS hijacking by middleboxes (see Section 3.1 and Appendix B) so responses carrying `A` records can only come from the censorship systems. By contrast, if resolved by domestic instances, root queries do not pass the censorship systems and should receive normal responses that carry delegation data of Top-Level Domains, as shown in Figure 1 (bottom).



```
$ dig @a.root-servers.net. censored-domain.com
;; ANSWER SECTION:
censored-domain.com.    120    IN     A     (CENSORED_ADDRESS)
```

```
$ dig @a.root-servers.net. censored-domain.com
;; ANSWER SECTION:
com.             172800   IN      NS      a.gtld-servers.net.
com.             172800   IN      NS      b.gtld-servers.net.
…
```

Fig. 1. Censored (top) and normal (bottom) responses to root queries

Leveraging this side channel provided by DNS censorship, Figure 2 overviews our approach to determining whether root queries are responded by domestic instances. From VPs in the Chinese mainland we send DNS queries of censored domains (e.g., `[nonce].google.com`) to each root server. Domains in the queries are prefixed with a nonce value, such that they must arrive at the root servers instead of being answered from cache (e.g., of middleboxes). If censored responses (Figure 1 top) are returned, the root queries must have passed the international gateway for instances overseas. Otherwise, if normal responses (Figure 1 bottom) are captured, the root queries are resolved by domestic instances.

Delay to root servers. To study the practical impact of domestic instances on root query performance, we also measure the Round-Trip-Time (RTT) of DNS queries from each VP to all 13 root servers. Because root servers are non-recursive (i.e., they never query other servers), the RTT can be acquired by simply sending DNS queries to root servers and recording the arrival time of their responses. Again, to ensure that the queries must arrive at the root servers, we register a

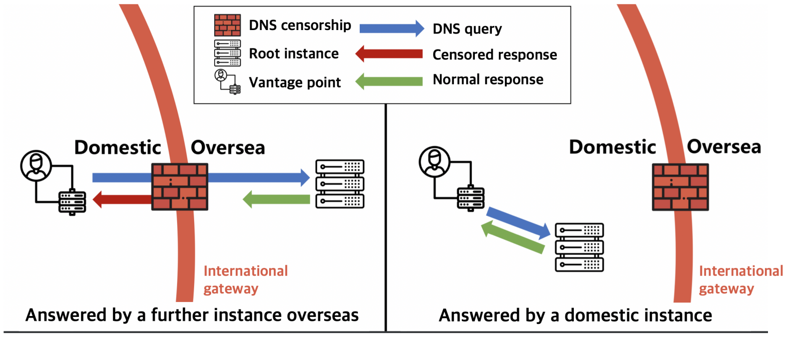


Fig. 2. Using DNS censorship to determine if domestic instances serve root queries

(non-censored) domain name exclusively for this task and prefix it with a nonce value in each query (i.e., `[nonce].example.com`).

4 Catchment Area of Domestic Root Instances

The Chinese mainland hosts 16 root server instances. However, as information about their hosting networks and peering ASes is not made public, it is still unclear which domestic networks can be served by them. In this section, we present our measurement results on the catchment area of domestic root server instances. We believe the results are helpful for future deployment of domestic instances to cover networks that are currently not served.

4.1 Which networks are served by domestic root instances?

Our experiment that leverages DNS censorship to measure catchment area started on Dec 5, 2020 and lasted for 72 hours. On each VP we send 10 DNS queries of *censored* domains (e.g., `[nonce].google.com`) to all 13 root servers every 24 hours and inspect whether they yield censored or normal responses. We retry if a query fails and the test issues 130 root queries per day from each VP.

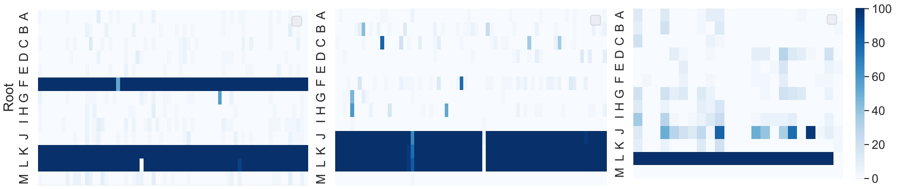
Table 3 shows the nation-wide ratio of root queries that receive normal responses (i.e., served by domestic instances) per ISP network. We first find that *all* root queries from CERNET VPs are answered domestically, which is expected because an unauthorized root server has been confirmed to exclusively serve CERNET [27]. We also find that the domestic instances of `F`, `J`, `K` and `L-Root` (all of them deploy instances in the Chinese mainland, as listed in Table 1) have *nation-wide catchment area* for VPs in *at least one of the three major ISPs*, as they answer over 95% of root queries from these networks (marked bold in Table 3, e.g., Telecom to `F-Root`). Zooming into individual VPs, as shown by Figure 3, we find that the catchment area of most domestic instances in a given ISP network only have minor differences between geo-locations.

For root servers that *do not* deploy domestic instances (e.g., `A` and `B-Root`, marked by darker backgrounds in Table 3), we do not expect that root queries

Table 3. Ratio of root queries that receive normal responses. Root servers with domestic instances (F, I, J, K and L) are marked with lighter backgrounds.

Root	Telecom	Unicom	Mobile	CERNET
A-Root	0.80%	0.50%	2.36%	**100.00%**
B-Root	0.94%	2.79%	2.08%	**100.00%**
C-Root	1.22%	3.29%	1.25%	**100.00%**
D-Root	0.85%	0.95%	5.69%	**100.00%**
E-Root	0.70%	0.00%	1.81%	**100.00%**
F-Root	**99.34%**	3.24%	1.53%	**100.00%**
G-Root	1.78%	0.99%	6.11%	**100.00%**
H-Root	0.85%	2.43%	2.22%	**100.00%**
I-Root	1.69%	0.09%	6.81%	**100.00%**
J-Root	1.27%	**98.24%**	23.47%	**100.00%**
K-Root	**100.00%**	**98.29%**	2.22%	**100.00%**
L-Root	**98.50%**	**98.38%**	**95.83%**	**100.00%**
M-Root	2.02%	0.05%	0.00%	**100.00%**
Total	23.84%	23.79%	11.65%	**100.00%**

trigger normal responses because they should arrive at instances beyond the international gateway and pass the censorship systems. However, in our results their ratio of normal responses does not reach 0%. We suppose that the normal responses are due to occasional failure of DNS censorship and potential unauthorized root servers deployed in domestic ISP networks, and we will discuss them in Section 6.



(a) Telecom (b) Unicom (c) Mobile

Fig. 3. Ratio of root queries that receive normal responses per VP. Darker cells indicate that more DNS queries from the corresponding VPs are resolved by domestic instances.

4.2 Why are some networks not served by domestic instances?

From Table 3 we also find that from some ISP networks to root servers that *do* deploy domestic instances, their ratio of getting normal responses is still very low, e.g., Mobile to K-Root, Telecom to J-Root and all three major ISPs to I-Root. The above results suggest that, these networks are still not well-served by the domestic instances.

The domestic I-Root instance. As shown in Table 1, one Global I-Root instance is deployed in the Chinese mainland, but it does not seem to be serving VPs in all three major ISPs. To locate the hosting network of this instance, we perform ICMP traceroute from 15 other controlled domestic VPs and find the second-last hops (the last hop is the root anycast address) belong to CSTNET (China Science and Technology Network). CSTNET offers Internet services to research institutions and hi-tech enterprises, but has a smaller user base than CERNET.

We then tried to figure out if CSTNET served by I-Root instance. As the looking glass platform does not cover CSTNET, we employ seven volunteer VPs in CSTNET and run the same experiments described in Section 3.2. Similarly, 100% queries receive censored responses when they carry sensitive domains, meaning that they pass the international gateway. However, the CSTNET VPs show an average delay of only 3.62ms to I-Root, which is significantly lower than three major ISPs (over 100ms on average, see results in Figure 4 of Section 5.1). Further, we ask all CSTNET VPs to send NSID queries [10] to I-Root. The responses carry an "s1.bei" string, and we confirm with Netnod (the operator of I-Root) that it represents the Beijing instance.

As a result, we conclude that the domestic I-root instance serve CSTNET only (possibly because three major ISPs do not peer with CSTNET or Netnod), and that it locates physically in the Chinese mainland but out of the international network gateway. In fact, this can be a result of a security incident in 2010 where this instance returned incorrect responses due to DNS censorship [11].

Other unshared domestic instances. Similarly, we locate the hosting networks of other domestic instances through ICMP traceroute. From Table 3, VPs in China Mobile are not served by domestic F, J and K-Root instances, and our traceroute shows that these instances locate in networks of China Telecom and China Unicom. VPs in China Telecom cannot access domestic J-Root instances, which we find in China Unicom.

Combined with discussions with ISPs and DNS root operators, we conclude that root instances in the Chinese mainland are typically advertised from ISP networks instead of IXes, and that root instances hosted in one ISP are typically *unshared* with other networks due to limits of BGP routing policies. For F and K-Root, some are *Local instances* (see Table 1) thus their catchment area does not cover networks of other ISPs. Meanwhile, major ISPs do not directly peer with each other[2], thus *Global instances* deployed in the Chinese mainland are not accessible from some domestic ISPs either (e.g., Mobile to J-Root).

5 Impact of Domestic Instances on Root Server Selection

In Section 4, we measure the catchment area of domestic instances by sending DNS queries directly to each root server. However, whether the instances can actually absorb root queries depend on how *recursive resolvers* in their catchment area select from 13 root servers (i.e., the domestic instances are queried only when the corresponding root servers are selected by recursive resolvers). In this section, we first study how deploying domestic instances affects the nation-wide

[2] We also tried to validate through inspecting BGP routing information in RouteViews [45]. However, the dataset has little coverage of ASes in China.

delay to root servers, and then show how it affects root server selection from the perspective of mainstream recursive resolver implementations.

5.1 Do domestic instances serve root queries with lower delay?

During the same time period and using the same set of VPs, we measure their delay to all 13 root servers. On each VP we perform 10 DNS queries of (custom) *non-censored* domains (e.g., `[nonce].example.com`) to each root server in every 24 hours and record their RTTs. We retry if a query fails and the test issues 130 root queries per day from each VP.

Figure 4 shows the delay from VPs in each ISP to 13 root servers. All root queries from CERNET are answered within 30ms because of an unauthorized root server. For the other three major ISPs, we confirm that *root server instances deployed in the Chinese mainland effectively serve networks within their catchment area with lower delay.* Corresponding Figure 4 with Table 3, we find that for networks with high ratio (>95%, the ISPs are also framed in Figure 4) of queries that are resolved by domestic instances, their delay to the corresponding root servers is significantly lower (50ms on average). By contrast, VPs spend hundreds of milliseconds to query instances beyond the international gateway when their network is not served by the domestic instances (e.g., Mobile to `F-Root`, Telecom to `J-Root` and all three major ISPs to `I-Root`), or when no instances are deployed in the Chinese mainland (e.g., `A` and `B-Root`).

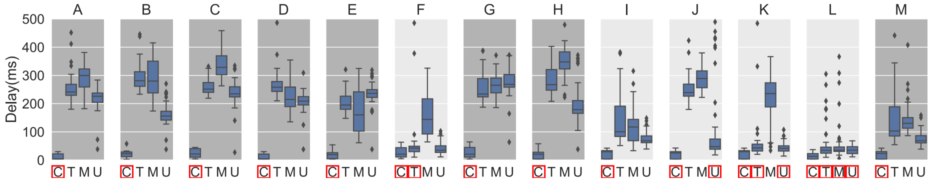


Fig. 4. Root server delay from VPs of different ISPs (C: CERNET, T: Telecom, M: Mobile, U: Unicom). Root servers with domestic instances are marked with lighter backgrounds. Framed networks correspond to high ratio (>95%) of domestically resolved queries in Table 3.

5.2 How do domestic instances affect root server selection?

Deploying domestic instances improves root query performance within their catchment area, and we further investigate *whether they will actually be selected* (among all 13 root servers) by recursive resolvers because of lowered delay. A series of previous study demonstrated that most resolver prefer authoritative servers with lowest latency. However, the actual reasons behind remain unrevealed. To this end, we take the perspective of mainstream recursive resolver implementations and review their source code of root selection algorithms.

We select the latest versions (as of Oct 2021) of four popular open-source recursive DNS implementations: BIND 9 [25] (9.17.18), Unbound [40] (1.13.2), Knot Resolver [16] (5.4.2) and PowerDNS Recursor [41] (4.5.6). To review and dynamically debug their root selection algorithms, we start a docker container [3] that installs Ubuntu 18.04 and links to a GDB remote debugger [4]. All DNS software is compiled from source code and started in the container. In the GDB debugger, we extract and review the root selection algorithms by following their execution.

Based on the result of source code reviewing and dynamic debugging, we also *quantify* the differences of root selection algorithms in a test network environment. We simulate and inspect 100,000 outgoing root queries from each DNS implementation. Our set of 13 root servers is: `RTT=10ms` (1 server), `RTT=50ms` (3 servers), `RTT=100ms` (5 servers), `RTT=250ms` (3 servers) and `RTT=500ms` (1 server). We do not consider response errors or timeouts during simulation.

Root server selection algorithm overview. We first find that all four recursive DNS implementations *reuse authoritative server (NS) selection algorithms* to select from root servers[3]. For this task, DNS standards [2] vaguely suggest that recursive resolvers should "find the best server to ask". From the comments in their source code, we find that the designers of DNS software reach a consensus that *least-latent servers will be preferred* (e.g., "Find best RTT in the bunch" in the comments of Unbound and "Address with smaller expected timeout is better" in the comments of Knot Resolver). However, as we will show later, the root selection results for Unbound and PowerDNS Recursor do not match this design goal due to untuned timing implementations.

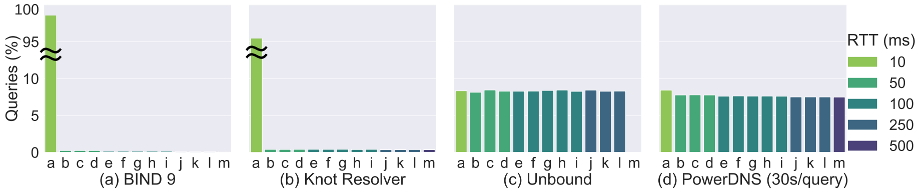
BIND 9 and Knot Resolver. BIND and Knot Resolver significantly prefer root servers with the smallest Round-Trip-Time (RTT) while trying other servers in fewer cases. Through source code debugging, our findings echo with [8, 50, 38] that use pure traffic analysis to demonstrate that resolvers prefer least-latent NSes. As shown in Figures 5(a) and (b), during our simulation BIND 9 selects the least-latent root server (`RTT=10ms`) in 98.1% cases, while the ratio for Knot Resolver is 95.3%.

Unbound and PowerDNS Recursor. Despite that they are designed to consider server RTT, Unbound and PowerDNS Recursor tend to select root servers *randomly* due to untuned timing implementations. Unbound selects from a set of servers randomly and disregards one only if its RTT is *400ms longer* than the least-latent. However, previous works [32, 23] and Figure 4 already show that only few locations across the globe witness a delay to root servers of over 400ms, thus Unbound will not remove any root server from consideration. As shown in Figure 5(c), during our simulation Unbound only disregards one root server (`RTT=500ms`) and evenly distributes root queries among all other servers.

For PowerDNS Recursor, it is designed to select the least-latent server and update its RTT, but *decays* the adjusted RTT for all servers with the same factor. The longer the query interval, the lower the decay factor and adjusted RTTs will be. Since root servers are not frequently queried ([12] shows the root query interval from 97% of addresses is longer than 100 seconds), the priority for

[3] As part of our contribution, we list pseudo-code and details of all root selection algorithms at `https://github.com/anonymous-researcher123/software-analysis/blob/main/software_analysis.pdf`

root servers not selected in the current round will be significantly increased in the next round. During our simulation, when we set the query rate to 30s/query, the selection already looks random (Figure 5(d)) because of significant decay. We believe the implementations of Unbound and PowerDNS do not properly switch to lower latency authoritative servers. And we will contact the developers of these two resolvers, and hope to improve the implementation in the future.



(a) BIND 9 (b) Knot Resolver (c) Unbound (d) PowerDNS (30s/query)

Fig. 5. Simulation results of root server selection algorithms with 100,000 root queries. (From source code we learn that the algorithm of PowerDNS depends on query interval and we show the results under 30s/query.)

Summary. Through source code debugging and simulation, we confirm that BIND and Knot Resolver significantly prefers least-latent root servers (for over 95% of queries). Considering their large share among recursive resolvers (e.g., BIND is deployed on over 60% of recursive resolvers [29, 1]), we conclude that domestic root server instances can effectively absorb queries from networks within their catchment area because of lowered delay compared to other root servers.

6 Discussion

Ethics. The major ethical concern of this study is sending DNS queries through VPs in the Chinese mainland. To acquire VPs, we leverage a commercial network looking glass platform and pay for their service. On each VP we send plain DNS queries over UDP to root servers, which is within the business scope of the platform. We also strictly limit the DNS query rate on each VP (at around 30 queries per hour) to comply with its service regulations.

To measure the catchment area of domestic root server instances, we leverage the DNS censorship mechanisms of China, which has been thoroughly studied by a series of previous works [9, 15, 31, 35]. In our methodology we use its known characteristics (i.e., injection of DNS responses) and do not study or provide new insights into the censorship systems. All domains carried by our queries are non-existent sub-domains (because of nonce prefixes) under benign Second-Level Domains (e.g., `[nonce].google.com` and our own domain name). We do not make connections to any censored IP address. Because we only perform DNS lookups, our study poses no harm or potential judicial risks to the VP operators.

Errors caused by DNS censorship failure. In Table 3 we find that for queries to root servers that do not deploy domestic instances (e.g., 3 major ISP networks to `A` and `B-Root`), a small portion (1% to 3%) receive uncensored normal responses. We confirm through a separate experiment that for the selected

VPs, the DNS censorship has an overall success rate of around 97% (details of the experiment are provided in Appendix A). As a result, the root queries do leave the international gateway and are resolved by instances overseas, but receive normal responses because of censorship failure.

Potential unauthorized root servers. From Table 3 we also find that, from some ISP networks to root servers without domestic instances (e.g., Mobile to `D-Root` and `G-Root`), the ratio of normal responses (5% to 6%) is higher than the average failure rate of DNS censorship (around 3%). In Figure 3 we zoom into the ratio of censored responses for each VP. A small number of VPs in China Mobile and China Unicom even receive around 20% to 40% uncensored normal responses from root servers without domestic instances, which is not likely a result of DNS censorship failure. It is also not caused by DNS hijacking or caching by middleboxes because we have removed such VPs (see Section 3.1) and use nonce domain prefixes to make every query unique. Finally, we suppose these normal responses are provided domestically by unauthorized root servers.

Recommendations. Our study reveals several major ISP networks that are not served by domestic root server instances and we make the following recommendations. 1) For networks that are not covered by the catchment area of nearby instances, if allowed in terms of political and commercial interests, we recommend BGP peering with the root server networks. Alternative measures that improve access to the RSS (e.g., running a local root copy [30]) can be adopted. 2) Root server operators may take these areas into prior consideration for future deployment of instances. 3) To inform operators about whether their networks can be served, we recommend making BGP peering information between ISPs and root servers more transparent (e.g., disclosing which networks host a root server or peer with them). 4) We do not recommend establishing or using unauthorized root servers which may cause security risks. 5) For developers of recursive DNS software, we recommend they review whether the implementation is consistent with their original design goals. 6) For the DNS community, while the status of root servers has been heavily monitored, systems that measure root servers from resolvers' perspective still need to be developed.

7 Conclusion

In this paper we study the practical effect of root server instances deployed in the Chinese mainland. Through design of novel methodology we measure the catchment area of domestic instances, and find that some of them are not accessible from major ISP networks due to limits of BGP routing policies. We also evaluate the impact of deploying new instances on root server selection in recursive resolvers by measuring root delay and reviewing source code of common recursive resolver implementations. While most software is designed to prefer less-latent servers, some do not meet this goal due to untuned implementations. Our results also show that domestic root instances significantly lower query delay from major ISP networks, which increases their possibility to absorb DNS queries. We believe that multiple parties should take actions to improve the stability and operational status of the DNS Root Server System in China.

Acknowledgement

We thank all the anonymous reviewers for insightful comments and valuable suggestions to improve this paper. This paper benefits from the contributions of a number of people, in particular, Ziqian Liu, provided excellent suggestions that improved the paper. This work was supported by the National Natural Science Foundation of China (grants 62102218, U1836213, U19B2034). Baojun Liu is partially supported by the Shuimu Tsinghua Scholar Program. Baojun Liu and Ying Liu are corresponding authors.

References

1. BIND DNS Holds Lead, https://www.serverwatch.com/server-news/bind-dns-holds-lead/
2. Domain Names - Concepts and Facilities. RFC 1034 (Nov 1987). https://doi.org/10.17487/RFC1034, https://rfc-editor.org/rfc/rfc1034.txt
3. Docker: Empowering App Development for Developers (2021), https://www.docker.com/
4. GDB: The GNU Project Debugger - GNU.org (2021), https://www.gnu.org/software/gdb/
5. Measuring the Health of the Domain Name System (2021), https://www.icann.org/en/system/files/files/dns-ssr-symposium-report-1-03feb10-en.pdf
6. Root Server Technical Operations Association (2021), https://root-servers.org/
7. Abley, J.: Hierarchical Anycast for Global Service Distribution (2003)
8. Ager, B., Mühlbauer, W., Smaragdakis, G., Uhlig, S.: Comparing DNS Resolvers in the Wild. In: Proceedings of the 10th ACM SIGCOMM conference on Internet measurement. pp. 15–21 (2010)
9. Anonymous: Towards a Comprehensive Picture of the Great Firewall's DNS Censorship. In: 4th USENIX Workshop on Free and Open Communications on the Internet (FOCI 14). USENIX Association, San Diego, CA (Aug 2014), https://www.usenix.org/conference/foci14/workshop-program/presentation/anonymous
10. Austein, R.: DNS Name Server Identifier (NSID) Option. RFC 5001 (Aug 2007). https://doi.org/10.17487/RFC5001, https://rfc-editor.org/rfc/rfc5001.txt
11. van Beijnum, I.: Misadventure at root I in China (2010), https://web.archive.org/web/20110622092029/http://arstechnica.com/tech-policy/news/2010/03/china-censorship-leaks-outside-great-firewall-via-root-server.ars
12. Castro, S., Wessels, D., Fomenkov, M., Claffy, K.: A Day at the Root of the Internet. ACM SIGCOMM Computer Communication Review **38**(5), 41–46 (2008)
13. Conrad, D.: Brief Overview of the Root Server System (2020), https://www.icann.org/en/system/files/files/octo-010-06may20-en.pdf
14. Conrad, D.R., Woolf, S.: Requirements for a Mechanism Identifying a Name Server Instance. RFC 4892 (Jun 2007). https://doi.org/10.17487/RFC4892, https://rfc-editor.org/rfc/rfc4892.txt
15. Crandall, J.R., Zinn, D., Byrd, M., Barr, E.T., East, R.: ConceptDoppler: A Weather Tracker for Internet Censorship. In: ACM Conference on Computer and Communications Security. pp. 352–365 (2007)
16. CZ.NIC: Knot Resolver (2021), https://www.knot-resolver.cz/

17. Fan, X., Heidemann, J., Govindan, R.: Evaluating Anycast in the Domain Name System. In: 2013 Proceedings IEEE INFOCOM. pp. 1681–1689. IEEE (2013)
18. Hardie, T.: Distributing Authoritative Name Servers via Shared Unicast Addresses. RFC 3258 (Apr 2002). https://doi.org/10.17487/RFC3258, `https://rfc-editor.org/rfc/rfc3258.txt`
19. IANA: Root Zone Management (2021), `https://www.iana.org/domains/root`
20. ICANN: Hosting IMRS in Your Network (2019), `https://www.dns.icann.org/imrs/host/`
21. ICANN: IMRS Peering Information (2019), `https://www.dns.icann.org/imrs/peering/`
22. ICANN: The ITHI (Identifier Technologies Health Indicators) Project (2021), `https://ithi.research.icann.org/about.html`
23. ISC: Atlas Data Viewer (2021), `https://atlas-vis.isc.org/`
24. ISC: BGP Peering Network with F-Root (2021), `https://www.isc.org/froot-peering/`
25. ISC: BIND 9 (2021), `https://www.isc.org/bind/`
26. ISC: Host an F-Root Node (2021), `https://www.isc.org/froot-process/`
27. Jones, B., Feamster, N., Paxson, V., Weaver, N., Allman, M.: Detecting DNS Root Manipulation. In: International Conference on Passive and Active Network Measurement. pp. 276–288. Springer (2016)
28. Koch, T., Katz-Bassett, E., Heidemann, J., Calder, M., Ardi, C., Li, K.: Anycast In Context: A Tale of Two Systems. In: Proceedings of the 2021 ACM SIGCOMM 2021 Conference. pp. 398–417 (2021)
29. Kührer, M., Hupperich, T., Bushart, J., Rossow, C., Holz, T.: Going Wild: Large-scale Classification of Open DNS Resolvers. In: Proceedings of the 2015 Internet Measurement Conference. pp. 355–368 (2015)
30. Kumari, W.A., Hoffman, P.E.: Running a Root Server Local to a Resolver. RFC 8806 (Jun 2020). https://doi.org/10.17487/RFC8806, `https://rfc-editor.org/rfc/rfc8806.txt`
31. Levis, P.: The Collateral Damage of Internet Censorship by DNS Injection. ACM SIGCOMM CCR **42**(3), 10–1145 (2012)
32. Liang, J., Jiang, J., Duan, H., Li, K., Wu, J.: Measuring Query Latency of Top Level DNS Servers. In: International Conference on Passive and Active Network Measurement. pp. 145–154. Springer (2013)
33. Liu, B., Lu, C., Duan, H., Liu, Y., Li, Z., Hao, S., Yang, M.: Who is Answering my Queries: Understanding and Characterizing Interception of the DNS Resolution Path. In: 27th USENIX Security Symposium (USENIX Security 18). pp. 1113–1128 (2018)
34. Liu, Z., Huffaker, B., Fomenkov, M., Brownlee, N., et al.: Two Days in the Life of the DNS Anycast Root Servers. In: International Conference on Passive and Active Network Measurement. pp. 125–134. Springer (2007)
35. Lowe, G., Winters, P., Marcus, M.L.: The Great DNS Wall of China. MS, New York University **21**, 1 (2007)
36. MaxMind: IP Geolocation and Online Fraud Prevention (2021), `https://www.maxmind.com/en/home`
37. Moura, G.C.M., de Oliveira Schmidt, R., Heidemann, J.S., de Vries, W.B., Müller, M., Wei, L., Hesselman, C.: Anycast vs. DDoS: Evaluating the November 2015 Root DNS Event. In: Gill, P., Heidemann, J.S., Byers, J.W., Govindan, R. (eds.) Proceedings of the 2016 ACM on Internet Measurement Conference, IMC 2016, Santa Monica, CA, USA, November 14-16, 2016. pp. 255–270. ACM (2016), `http://dl.acm.org/citation.cfm?id=2987446`

38. Müller, M., Moura, G.C., de O. Schmidt, R., Heidemann, J.: Recursives in the Wild: Engineering Authoritative DNS Servers. In: Proceedings of the 2017 Internet Measurement Conference. pp. 489–495 (2017)
39. Netnod: Host an I-Root (2021), `https://www.netnod.se/host-an-i-root`
40. NLnet Labs: Unbound (2021), `https://www.nlnetlabs.nl/projects/unbound/about/`
41. PowerDNS: PowerDNS Recursor (2021), `https://www.powerdns.com/recursor.html`
42. RSSAC: RSSAC002: RSSAC Advisory on Measurements of the Root Server System (2015)
43. RSSAC: RSSAC057: Requirements for Measurements of the Local Perspective on the Root Server System (2021)
44. Staff, R.N.: Ripe Atlas: A Global Internet Measurement Network. Internet Protocol Journal **18**(3) (2015)
45. University of Oregon: Route Views Project (2021), `http://www.routeviews.org/routeviews/`
46. Weaver, N., Kreibich, C., Paxson, V.: Redirecting DNS for Ads and Profit. FOCI **2**, 2–3 (2011)
47. Weber, J.: Detect DNS Spoofing: dnstraceroute (2016), `https://weberblog.net/detect-dns-spoofing-dnstraceroute/`
48. Weinberg, Z., Cho, S., Christin, N., Sekar, V., Gill, P.: How to Catch when Proxies Lie: Verifying the Physical Locations of Network Proxies with Active Geolocation. In: Proceedings of the Internet Measurement Conference 2018, IMC 2018, Boston, MA, USA, October 31 - November 02, 2018. pp. 203–217. ACM (2018), `https://dl.acm.org/citation.cfm?id=3278551`
49. Wicaksana, M.: Ipv4 vs ipv6 anycast catchment: a root dns study (August 2016), `http://essay.utwente.nl/70921/`
50. Yu, Y., Wessels, D., Larson, M., Zhang, L.: Authority Server Selection in DNS Caching Resolvers. ACM SIGCOMM Computer Communication Review **42**(2), 80–86 (2012)

A Success Rate of DNS Censorship

We use the same set of VPs from the looking glass platform to measure the success rate of DNS censorship. On each VP we launch DNS queries of censored domain names (e.g., `[nonce].google.com`) to a self-built DNS server located in the US. This design ensures that all DNS queries should leave the international gateway and pass the DNS censorship system. As shown in Table 4, the DNS censorship system has an overall success rate of 97%, which explains why 1% to 3% of DNS queries to root servers without domestic instances receive (uncensored) normal responses (see Section 4.1).

To further confirm that the queries leave the international gateway, in Figure 6 we plot the RTTs of normal responses from `A-Root` (without domestic instances). All normal responses to VPs in China Telecom, Unicom and Mobile have an RTT of around 200ms, significantly longer than responses from CERNET (a baseline delay for domestic responses). From the results we are confident that the 1% to 3% normal responses come from root instances overseas, rather than domestic servers.

Table 4. Success rate of DNS censorship

ISP	Ratio of Censored Responses
China Telecom	96.25 %
China Unicom	96.56 %
China Mobile	97.82 %
CERNET	94.86 %
Total	**96.72 %**

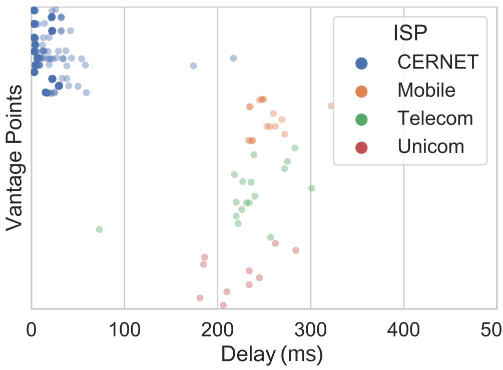


Fig. 6. Delays of normal responses from `A-Root` to domestic VPs

B Removed VPs that Perceive DNS Hijacking Accidents

During VP validation, we find multiple DNS hijacking accidents in large ISP networks. As shown in Table 5, domains are pointed to rogue addresses or show negative results (NXDOMAIN). Our results echo with [33] reporting that DNS hijacking behaviors are more prevalent for VPs of China Mobile. All VPs are then removed from consideration.

Table 5. Removed VPs that perceive DNS hijacking

Response Type	Telecom	Unicom	Mobile	CERNET	Total
Polluted IP	28	13	39	6	**86**
Localhost IP	2	0	8	2	**12**
NXDOMAIN	4	7	16	2	**29**
Empty	0	0	2	0	**2**
Other IP	5	3	3	0	**11**
Total	**39**	**23**	**68**	**10**	