



Improving Privacy for GeoIP DNS Traffic

Lanlan Pan¹(✉), Xuebiao Yuchi², Xin Zhang³, Anlei Hu³,
and Jian Wang¹

¹ Geely Automobile Research Institute, Zhejiang 315336, China
abbypan@gmail.com

² Chinese Academy of Sciences, Beijing 100190, China

³ China Internet Network Information Center, Beijing 100190, China

Abstract. Many authoritative nameservers today support GeoIP feature. EDNS Client Subnet (ECS) extension helps GeoIP authoritative nameserver to address the public recursive resolver's proximity IP problem. However, ECS raises some privacy concerns since recursive resolver leaks client subnet information on the resolution path to the authoritative nameserver. In this paper we introduce an EDNS ISP Location (EIL) extension, to make privacy improvement for GeoIP DNS traffic while preserve the ECS optimization on the end-user experience, reduce response latency, and increase cache-hit rate. We analysis 910.9K Chinese IPv4 CIDR/24 subnets, find that 479.9K TEL subnets, 234.0K UNI subnets, and 66.3K MOB subnets can enable EIL to optimize DNS traffic.

Keywords: DNS · Privacy · GeoIP · Client subnet · ECS · EIL

1 Introduction

In order to bring the web content as close to the users as possible, many authoritative nameservers support GeoIP feature, return different responses based on the perceived geographical location of the resolvers' IP addresses [1–6].

As Fig. 1 shows, there are two critical factors that can affect the response accuracy of authoritative nameserver:

- (1) Proximity IP Problem: Is the resolver's IP address close enough to the client's IP address?
- (2) GeoIP Database Problem: Does the authoritative nameserver use an GeoIP database with high quality?

Public recursive resolvers such as Google Public DNS and OpenDNS offer free DNS resolution services for global users. These servers are not close enough to many users since the public recursive service providers couldn't deploy servers among each country and each ISP's network [7].

Therefore, public recursive resolvers face to serious proximity IP problem. To counter this problem, Google proposes an EDNS Client Subnet (ECS) extension [8] to

This is an extended version of an earlier extended abstract presented at the International Conference on Privacy, Security and Trust, 2018.

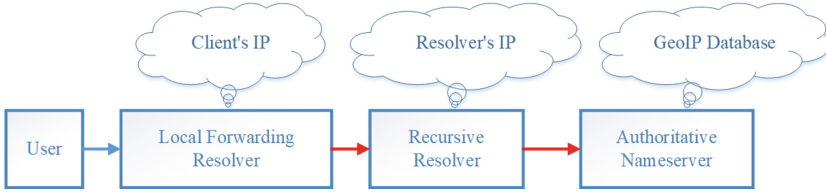


Fig. 1. GeoIP DNS traffic.

carry part of the client’s IP address in the DNS packets for authoritative nameserver. As Fig. 2 shows, authoritative nameserver can directly use client subnet information to better understand where the end user is, while ignoring the resolver’s IP address.

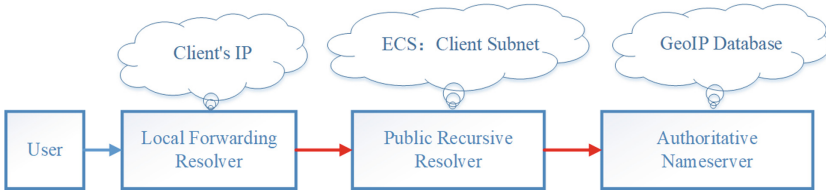


Fig. 2. ECS extension worked on GeoIP DNS traffic.

However, ECS also raises some privacy concerns because it leaks client subnet information on the resolution path to the authoritative nameserver. In [9], Kintis pointed out that ECS makes DNS communications less private: the potential for mass surveillance is greater, and stealthy, highly targeted DNS poisoning attacks become possible. Doileak [10] describe the privacy risk of ECS and why using a public DNS server might not improve your privacy.

To find the right balance between privacy improvement and end-user experience optimization, in this paper we introduce an EDNS ISP Location (EIL) extension. Compared with ECS, EIL can counter the proximity IP problem and GeoIP database problem more effectively, and improve privacy for GeoIP DNS Traffic.

The remainder of this paper is organized as follows. In Sect. 2, we discuss some related DNS privacy protection technologies. In Sect. 3, we describe the EIL extension in detail. From Sect. 4 to Sect. 6, we discuss response accuracy enhancement, privacy improvement, and operational benefit of EIL. In Sect. 7, we show our experiment on some GeoIP domains, and analyze the EIL effect. Finally, in Sect. 8, we discuss our work and conclude the paper.

2 DNS Privacy Protection Technologies

As Fig. 3 shows, most DNS privacy protection technologies [11, 12] can be divided into two groups. However, existing technologies are hard to provide user privacy controls on recursive resolvers that support ECS.

- Encrypting DNS Traffic

DNS over TLS [13], DNSCurve [14], DNSCrypt [15] and Confidential DNS [16] are different technologies to encrypt DNS traffic, they can improve the privacy on the resolution path, while none of them has any influence on the nameservers.

- Reducing Information Leakage to DNS Server

Root loopback [17] and QNAME minimization [18] can hide domain query information from root and TLD, while they are not designed for Client's IP privacy.

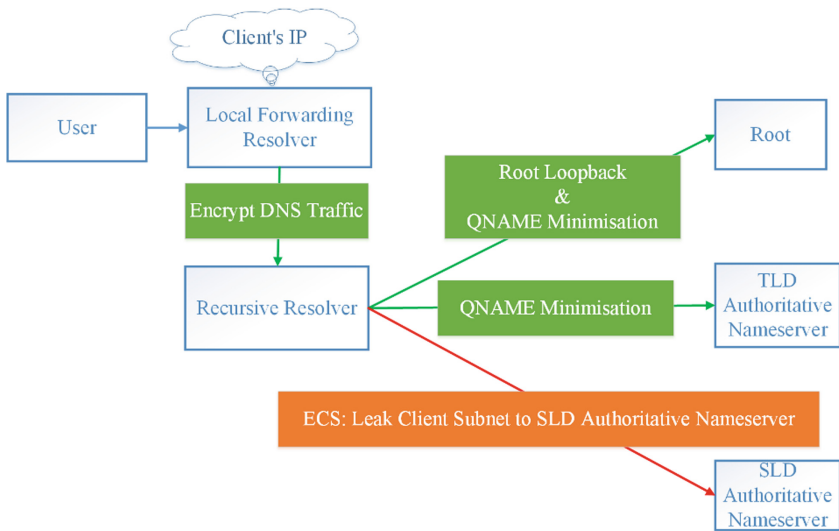


Fig. 3. DNS privacy protection technologies.

3 EDNS ISP Location (EIL) Extension

The EDNS ISP Location (EIL) extension proposed in this paper is similar to ECS. However, EIL includes the GeoIP information of client's IP in DNS packets, not client subnet information. Authoritative nameserver could provide a better answer by using GeoIP information of client's IP in EIL.

EIL can be added in DNS queries sent by recursive resolvers or local forwarding resolvers in a way that is transparent to stub resolvers and end users. EIL is only defined for the Internet (IN) DNS class.

3.1 Structure

EIL is structured as follows (Fig. 4):

- **OPTION-CODE**, 2 octets, defined in RFC6891 [19]. EDNS option code should be assigned by the IANA.

- **OPTION-LENGTH**, 2 octets, defined in RFC6891, contains the length of the payload (everything after OPTION-LENGTH) in octets.
- **COUNTRY**, 2 octets, uppercase, defined in ISO3166 [20], indicates the country information of the client’s IP. For example, The COUNTRY of China is CN.
- **AREA**, 6 octets, uppercase, defined in ISO3166 country subdivision code, indicates the area information of the client’s IP. For example, The AREA of Fujian Province in China is 35.
- **ISP**, 4 octets, uppercase, indicates the ISP information of the client’s IP, using shortcut names. ISP shortcut names are unique within the context of the COUNTRY. As Table 1 shows, the shortcut name of China Telecommunications Corporation is TEL.

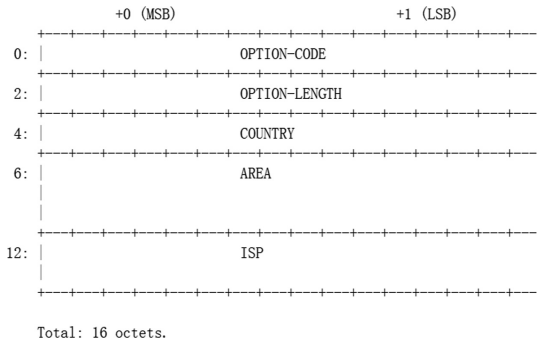


Fig. 4. EIL structure.

All fields of EIL are in network byte order. We use short names in the fields to limit the data size of EIL, decrease the DDoS risk. The null value 0×20 signifies that the field is unknown. If all fields in EIL are set to null value, means that client doesn’t want to use EIL.

Table 1. China ISP.

ISP	ISP fullname
TEL	China Telecommunications Corporation
UNI	China United Network Communications
MOB	China Mobile Communications Corporation
TIE	China Tietong Telecommunications Corporation
EDU	China Education and Research Network

3.2 GeoIP Information

As Fig. 5 shows, Maxmind [21] gives the GeoIP information:

- Location: Quanzhou, Fujian, China, Asia
- ISP name: China Telecom

We can map Client's IP 61.154.123.91 into EIL <CN, 35, TEL>. Compared to ECS's client subnet such as 61.154.123.0/20, EIL contains very few sensitive information because it is associated with a very broad geographic area.

GeoIP2 City Database Demo

IP Addresses

Enter up to 25 IP addresses separated by spaces or commas. You can also [test you](#)

GeoIP2 City Results

IP Address	Country Code	Location	Postal Code	Approximate Coordinates*	Accuracy Radius	ISP
61.154.123.91	CN	Quanzhou, Fujian, China, Asia		24.9139, 118.5858	100	China Telecom

Fig. 5. Maxmind GeoIP information.

3.3 Deploy

Take Fig. 6 for example, when a public recursive resolver receives a DNS query from local forwarding resolver, it can map the client's IP to EIL <COUNTRY, AREA, ISP> information, then send EIL query to the authoritative nameserver. Using the GeoIP information specified in the EIL of DNS query, the authoritative nameserver can generate a tailored response.

Compared with ECS, EIL will move the GeoIP information mapping work from authoritative nameserver to recursive resolver, lighten the burden of authoritative nameserver, while it will increase DDoS risk on recursive resolver.

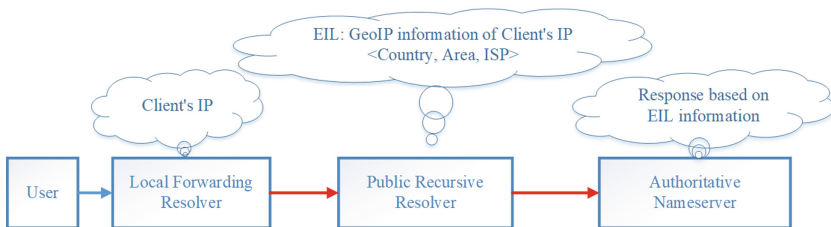


Fig. 6. EIL extension worked on GeoIP DNS traffic.

4 Response Accuracy Enhancement

4.1 Proximity IP Problem

ECS solves the proximity IP problem by generating the client subnet information from client's IP address.

Similar to ECS, EIL's GeoIP information <COUNTRY, AREA, ISP> is generated from client's IP address. Therefore, EIL also solves the proximity IP problem to GeoIP-enabled authoritative nameserver.

4.2 GeoIP Database Problem

GeoIP database quality affect response accuracy heavily. In ECS traffic mode, different GeoIP-enabled authoritative nameservers probably build up different GeoIP databases for their tailor response. However, it is very difficult to ensure huge amounts of authoritative nameservers update their own GeoIP databases timely.

On the other hand, public recursive resolvers such as GoogleDNS and OpenDNS serve magnanimity clients in global, they are far more probably to build up high quality GeoIP database than many small authoritative nameservers. Therefore, if public recursive resolvers such as GoogleDNS and OpenDNS support EIL, they can make sure huge amounts of authoritative nameservers return tailored response based on more precious GeoIP information, globally synchronized the enhancement of authoritative nameservers' response accuracy.

5 Privacy Improvement

5.1 Mitigating Client Subnet Leakage

The biggest privacy concern on ECS is that client subnet information is personally identifiable.

The more domains publish their zones on a third-party GeoIP-enabled authoritative nameserver, the more end user privacy information can be gathered by the third-party authoritative nameserver according to the ECS queries. Moreover, many authoritative nameservers only accept plaintext DNS queries, which means that the client subnet information is transparent on the resolution path from recursive resolver to authoritative nameserver.

EIL replaces the sensitive client subnet information to aerial view GeoIP information for user privacy protection. The GeoIP information is generated from Client's IP, not from user's physical geolocation. Even with EIL's most precise GeoIP information, authoritative nameserver can't identify sensitive personal information, and not any sensitive personal information is in plaintext DNS traffic from recursive resolver to authoritative nameserver. That is, EIL improves user privacy by sending less personal sensitive data than ECS.

5.2 Combating Targeted Censorship

DNS traffic is in plaintext by default. It is easily to be blocked or poisoned on internet. On plaintext mode, ECS query is fragile to targeted client subnet censorship.

However, since EIL's GeoIP information covers much bigger area than ECS's client subnet information, EIL will be stronger at monitoring targeted DNS censorship attack.

Encrypting the DNS traffic will be helpful to defense the targeted censorship in the future.

6 Operational Benefit

6.1 Cache-Hit Rate of Recursive Resolver

ECS sends the query with client subnet, which means that recursive resolvers send a new query to authoritative nameservers for each client subnet, even when they have known the response for some other GeoIP-closed client subnets. In fact, thousands of client subnets usually visit only a few target servers, there are many redundancy queries which can cause adverse effect on the average of response latency of recursive resolvers. Figure 7 shows a sample of ECS redundancy queries for www.qq.com.

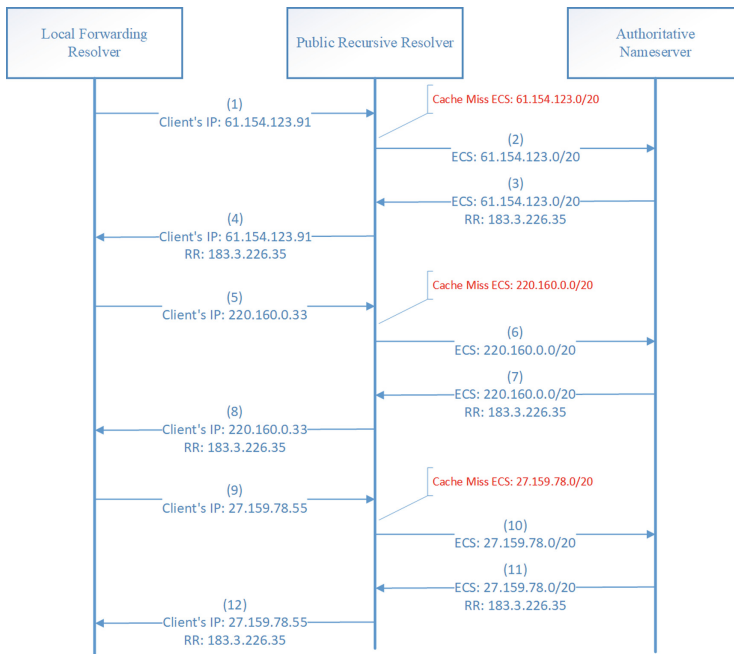


Fig. 7. ECS cache miss.

Each EIL GeoIP information covers huge amounts of client subnets. Therefore, compared to ECS redundancy queries to authoritative nameservers, EIL can sharply rise the cache-hit rate, reduce the response latency of recursive resolvers, lighten the burden of authoritative nameservers. Since EIL sends much less queries to authoritative nameservers, it can also improve privacy like qname minimization [18]. Figure 8 shows a sample of EIL optimized query for www.qq.com.

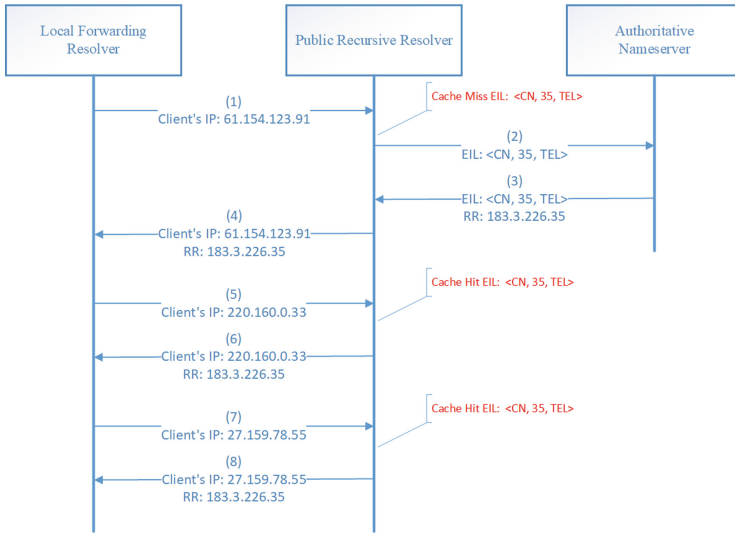


Fig. 8. EIL cache hit.

6.2 Cache Size of Recursive Resolver

EIL contains a whitelist for COUNTRY, AREA and ISP, which can be maintained by the IETF. Authoritative nameservers that support EIL can only response the EIL queries according to the whitelist. Recursive resolver that support EIL can only cache the EIL responses according to the whitelist too. Therefore, the EIL cache size of recursive resolver is related to the row count in the <COUNTRY, AREA, ISP> GeoIP information whitelist.

However, the ECS cache size of recursive resolver grows up with the number of client subnets. Obviously, under IPv6 environment, the EIL cache size will be much smaller than ECS.

Let's take the example of China in Table 2. There are 34 Areas in China. As Table 1 shows, TEL, UNI, MOB, TIE and EDU are the top 5 ISP in China. Consider about the null value of AREA and ISP, there will be 210 configurations on the authoritative nameserver to match the GeoIP information of China. This is the maximum cache size of EIL on recursive resolver for China.

Table 2. GeoIP information in China.

GEOIP TYPE	Configuration number
AREA + ISP	$34 * 5 = 170$
AREA + NULL ISP	$34 * 1 = 34$
NULL AREA + ISP	$1 * 5 = 5$
NULL AREA + NULL ISP	$1 * 1 = 1$
Total	$170 + 34 + 5 + 1 = 210$

6.3 DDoS Attack

To defend spoofed IP addresses, nameservers can optionally implement EIL query only when the query is from a TCP connection. In case of pseudo-random sub-domain attack, nameservers may encounter error EIL queries padding with some random error string. As we have limited the data size of EIL, the defense cost will be smaller than sub-domain attack.

In strict defense mode, authoritative nameserver can refuse all error EIL query for security. However, for a better user experience, recursive resolver can also make a better EIL query instead of refusing if it thinks itself can afford.

7 Experiment

We describe the data collection and analysis of our study, which try to show the EIL major improvements on GeoIP DNS traffic described in previous sections. Our experiment code can be found in Github [22].

7.1 Data Collection

As Table 3 shows, we totally collect 910.9K Chinese IPv4 CIDR/24 subnets for experiment, which cover top 3 China ISPs and 31 Areas.

Table 3. Number of IPv4 CIDR/24 subnets from China Top 3 ISPs.

Country	ISP	Number of IPv4 CIDR/24 subnets
CN	TEL	497.8K
CN	UNI	272.3K
CN	MOB	140.8K

For each subnet, we send the ECS query for www.qq.com to authoritative nameserver 123.151.66.83, get the tailored response, and add the GeoIP information.

Table 4 takes the client subnet 61.154.123.0/24 for example.

- We send the www.qq.com query to authoritative nameserver 123.151.66.83, with an ECS 61.154.123.0/24 option.

- 123.151.66.83 return the tailored response 183.3.226.35.
- We add the GeoIP information of response 183.3.226.35, which is <CN, 44, TEL>.
- We map the ECS 61.154.123.0/24 into EIL <CN, 35, TEL> GeoIP information.
- We finally build up the json-style response data for 61.154.123.0/24.

Table 4. Example ECS query and data collection.

Type	Value
Client subnet	61.154.123.0/24
ECS query	\$ dig +short @123.151.66.83 www.qq.com +subnet = 61.154.123.0/24 www.qq.com 183.3.226.35
Response data	{ domain: "www.qq.com", ecs_prefix: "61.154.123.0", ecs_mask: "24", eil_country: "CN", eil_area: "35", eil_isp: "TEL", response: [{ ip: "183.3.226.35", ip_country: "CN", ip_area: "44", ip_isp: "TEL" }] }

7.2 Analysis

We count the subnets each response IP covered by <eil_country, eil_area, eil_isp> GeoIP information.

Table 5 shows <CN, 44, TEL> for example, the 44 means Guangdong area. We totally check 81720 subnets from <CN, 44, TEL>. The most frequent response IP is 183.3.226.35, which covers 80785 subnets, 98.85585%. Obviously, 183.3.226.35 virtually monopolize the <CN, 44, TEL> subnets' response, and the authoritative nameserver of www.qq.com supports GeoIP feature. If the authoritative nameserver and recursive resolver both enable EIL, then recursive resolver can directly return 183.3.226.35 as response to <CN, 44, TEL> subnets, avoid meaningless redundant ECS query traffic.

Table 6 shows <CN, 11, UNI> for example, the 31 means Beijing area. We totally check 32930 subnets from <CN, 11, UNI>. The top 2 frequent response IPs are 125.39.52.26 and 180.163.26.39, they cover 32520 subnets, 98.75493%. If the authoritative nameserver and recursive resolver both enable EIL, then recursive resolver can directly return 125.39.52.26 and 180.163.26.39 as response to <CN, 11, UNI> subnets, avoid meaningless redundant ECS query traffic.

Table 5. Response for <CN, 44, TEL>.

ID	Country	Area	ISP	Response IP	Subnets	Percent	Accumulate subnets	Accumulate percent
1	CN	44	TEL	183.3.226.35	80785	98.85585%	80785	98.85585%
2	CN	44	TEL	121.51.142.21	408	0.499266%	81193	99.35512%
3	CN	44	TEL	180.163.26.39	332	0.406265%	81525	99.76138%
4	CN	44	TEL	58.250.137.36	119	0.145619%	81644	99.90700%
5	CN	44	TEL	123.151.137.18	59	0.072198%	81703	99.97920%
6	CN	44	TEL	125.39.52.26	16	0.019579%	81719	99.99878%
7	CN	44	TEL	61.129.7.47	1	0.001224%	81720	100%

Table 6. Response for <CN, 11, UNI>.

ID	Country	Area	ISP	Response IP	Subnets	Percent	Accumulate Subnets	Accumulate Percent
1	CN	11	UNI	125.39.52.26	27289	82.86972%	27289	82.86972%
2	CN	11	UNI	180.163.26.39	5231	15.88521%	32520	98.75493%
3	CN	11	UNI	182.254.50.164	213	0.646827%	32733	99.40176%
4	CN	11	UNI	123.151.137.18	154	0.467659%	32887	99.86942%
5	CN	11	UNI	58.247.214.47	18	0.054661%	32905	99.92408%
6	CN	11	UNI	58.250.137.36	16	0.048588%	32921	99.97267%
7	CN	11	UNI	61.129.7.47	4	0.012147%	32925	99.98482%
8	CN	11	UNI	121.51.36.46	3	0.00911%	32928	99.99393%
9	CN	11	UNI	183.3.226.35	2	0.006073%	32930	100%

Table 7 shows <CN, 11, MOB> for example, the 31 means Beijing area. We totally check 53508 subnets from <CN, 11, MOB>. The top 3 frequent response IPs are 111.30.132.101, 121.51.142.21 and 121.51.36.46, they cover 51451 subnets, 96.15572%.

If we set the EIL enable threshold of the authoritative nameserver is top 2 frequent response IPs' accumulate percent is not less than 98.5%, authoritative nameserver can disable EIL response for <CN, 11, MOB> subnets, and the recursive resolver can send old ECS query traffic as before.

Table 7. Response for <CN, 11, MOB>.

ID	Country	Area	ISP	Response IP	Subnets	Percent	Accumulate subnets	Accumulate percent
1	CN	11	MOB	111.30.132.101	36041	67.35628%	36041	67.35628%
2	CN	11	MOB	121.51.142.21	8093	15.12484%	44134	82.48112%
3	CN	11	MOB	121.51.36.46	7317	13.67459%	51451	96.15572%
4	CN	11	MOB	111.30.144.71	2057	3.844285%	53508	100%

Table 8. EIL enable threshold of the authoritative nameserver.

Pseudocode
<pre> for each <COUNTRY, AREA, ISP> { for my \$id (1 .. \$max_id){ if (Accumulate Percent >= \$min_percent){ enable EIL; set top \$id Response IPs as EIL response. } } } </pre>

Table 8 shows EIL enable threshold pseudocode of the authoritative nameserver.

For example, we can set $\$max_id = 2$ and $\$min_percent = 98.5\%$, Table 9 shows the EIL enable status. For TEL ISP, the authoritative nameserver of www.qq.com can enable EIL in 28 areas, which covered 479.9K subnets, 96.40245%. For UNI ISP, 26 areas can enable EIL, which covered 234.0K subnets, 85.92825%. For MOB ISP, 4 areas not enable EIL, the area codes are 11(Beijing), 32(Jiangsu), 31(Shanghai), 14 (Shanxi). We can find that responses for MOB ISP are not as steady as TEL ISP and UNI ISP, in this case, reserve ECS query can help for website traffic optimization.

Table 9. EIL enable decision for $\$max_id = 2$, $\$min_percent = 98.5\%$.

ISP	Enable EIL	Areas	Subnets	Percent
TEL	Yes	28	479.9K	96.40245%
	No	3	17.9K	3.597552%
UNI	Yes	26	234.0K	85.92825%
	No	5	38.3K	14.07175%
MOB	Yes	27	66.3K	47.10294%
	No	4	74.5K	52.89706%

8 Conclusion

We can't neglect the internet content delivery acceleration brought by ECS. The goal of EIL is to make privacy improvement for GeoIP DNS traffic while preserve the ECS optimization on the end-user experience, reduce response latency, and increase cache-hit rate.

We believe that EIL can provide user privacy controls both on public recursive resolvers and authoritative nameservers. Our future work is to do more experiments in China's network environment. We wish to apply the EIL into the real DNS traffic in the future, the IETF draft of EIL can be found in [23].

References

1. Amazon Route 53: Geolocation Routing. <http://docs.aws.amazon.com/Route53/latest/DeveloperGuide/routing-policy.html#routing-policy-geo>
2. Using the GeoIP Features in BIND 9.10. <https://kb.isc.org/article/AA-01149/0>
3. DYN Predefined Geographic Groups of Traffic Director. <https://help.dyn.com/traffic-director-predefined-geographic-regions/>
4. Gdnisd Plugin GeoiP. <https://github.com/gdnisd/gdnisd/wiki/GdnisdPluginGeoiP>
5. PowerDNS GeoIP backend. <https://doc.powerdns.com/md/authoritative/backend-geoiP/>
6. Microsoft Use DNS Policy for Geo-Location Based Traffic Management with Primary Servers. <https://docs.microsoft.com/en-us/windows-server/networking/dns/deploy/primary-geo-location>
7. Which CDNs support edns-client-subnet. <http://www.cdnplanet.com/blog/which-cdns-support-edns-client-subnet/>
8. Contavalli, C., van der Gaast, W., Lawrence, D., Kumari, W.: Client Subnet in DNS Queries. RFC7871 (2016)
9. Kintis, P., Nadji, Y., Dagon, D., Farrell, M., Antonakakis, M.: Understanding the privacy implications of ECS. In: Caballero, J., Zurutuza, U., Rodríguez, Ricardo J. (eds.) DIMVA 2016. LNCS, vol. 9721, pp. 343–353. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-40667-1_17
10. The privacy risk of edns-subnet-client (ECS). <https://www.doileak.com/blog-Public-DNS-might-not-%20improve-privacy.html>
11. Bortzmeyer, S.: DNS privacy considerations. RFC 7626 (2015)
12. Grothoff, C., Wachs, M., Ermert, M., Appelbaum, J.: NSA's MORECOWBELL: Knell for DNS
13. Hu, Z., et al.: Specification for DNS over Transport Layer Security (TLS). RFC 7858 (2016)
14. Dempsey, M.: Dnscurve: link-level security for the domain name system. Work in Progress, draft-dempsey-dnscurve-01 (2010)
15. DNSCrypt. <https://dnscrypt.org/>
16. Wijngaards, W., Wiley, G.: Confidential DNS. IETF Draft (2015). <https://tools.ietf.org/html/draft-wijngaards-dnsop-confidentialdns-03>
17. Kumari, W., Hoffman, P.: Decreasing Access Time to Root Servers by Running One on Loopback. RFC 7706 (2015)
18. Bortzmeyer, S.: DNS Query Name Minimisation to Improve Privacy. RFC7816 (2016)
19. Damas, J., Graff, M., Vixie, P.: Extension mechanisms for DNS (EDNS (0)). RFC 6891 (2013)
20. ISO 3166 Country Codes. http://www.iso.org/iso/country_codes
21. Maxmind GeoIP2 City Database. <https://www.maxmind.com/en/geoip-demo>
22. dns_test_eil. https://github.com/abbypan/dns_test_eil
23. Pan, L., Fu, Y.: ISP Location in DNS Queries. IETF Draft (2017). <https://datatracker.ietf.org/doc/draft-pan-dnsop-edns-location/>