EIL: Dealing with the Privacy Problem of ECS

Lanlan Pan, Xin Zhang, and Anlei Hu
China Internet Network Information Center, Beijing 100190, China
{panlanlan, zhangxin01, huanlei}@cnnic.cn

***Abstract*—Many authoritative servers today return different responses based on the perceived geographical location of the resolvers' IP addresses, to bring the content as close to the users as possible. RFC7871 proposes an EDNS Client Subnet (ECS) extension to carry part of the client's IP address in the DNS packets for authoritative server. Compared to the resolver's IP address in the DNS packets, ECS can help the authoritative server to guess the user's geographical location more precisely. However, ECS raises some privacy concerns because it leaks client subnet information on the resolution path to the authoritative server. This paper describes an EDNS ISP Location (EIL) extension to address the privacy problem of ECS, find the right balance between privacy improvement and end-user experience optimization. Moreover, EIL can reduce dependence on the IP geolocation database quality, which is crucial to DNS response accuracy in ECS.**

Keywords—DNS; recursive; authoritative; CDN; ECS; EIL

I. INTRODUCTION

Many authoritative servers today return different responses based on the perceived geographical location of the resolvers' IP addresses, to bring the content as close to the users as possible. As Figure 1 shows, there are two critical factors that affect the response accuracy of authoritative server:

1) Is the resolver's IP address close to the client's IP address?

2) Is the IP geolocation database used by the authoritative server with high quality?

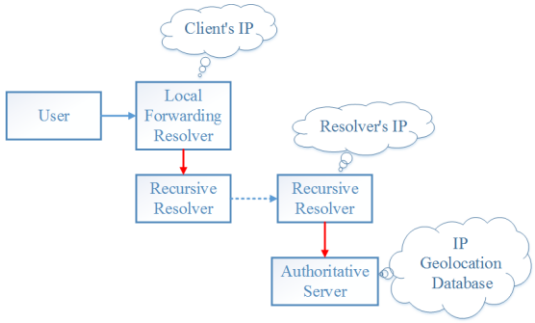

Figure 1. Overview of DNS Query

Public recursive resolvers such as Google Public DNS and OpenDNS offer free DNS resolution services for global users. But these servers are not close to many users because the public recursive service providers couldn't deploy servers in every country and every ISP's network[1]. To counter this problem, RFC7871[2] proposes an EDNS Client Subnet (ECS) extension to carry part of the client's IP address in the DNS packets for authoritative server. Authoritative server can directly use the client subnet information in ECS to better determine where the end user is, ignoring the resolver's IP address. Users can get big benefits from ECS, especially when they visit latency-sensitive CDN websites.

However, ECS also raises some privacy concerns because it leaks client subnet information on the resolution path to the authoritative server. In [3], 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.

This paper describes an EDNS ISP Location (EIL) extension to address the privacy problem of ECS, finds the right balance between privacy improvement and end-user experience optimization. The remainder of this paper is organized as follows. In Section II, we discuss some related DNS privacy protection technologies. In Section III, we describe the EIL extension in detail and compare EIL to ECS. In Section IV, we discuss some privacy and security concerns on EIL. Finally, in Section V, we discuss our work and conclude the paper.

II. ECS PRIVACY LEAKAGE

As Figure 2 shows, if the recursive resolver supports ECS, the end user's privacy leakage comes. Client subnet information is sent to authoritative servers transparently.

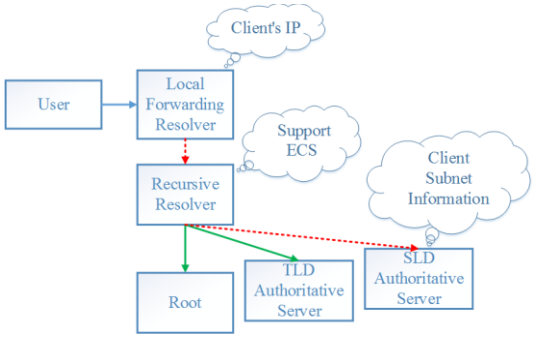

Figure 2. Client Subnet Information Leakage.

Bortzmeyer described the privacy issues associated with the use of the DNS by Internet users in RFC7626[4]. Grothoff discussed many proposed DNS privacy protection technologies in [5]. In [4] and [5], we can see that most of DNS privacy protection technologies can be divided into two groups:

1) Encrypt DNS Traffic
2) Reduce Information Leakage to DNS Server

However, these technologies are hard to provide user privacy controls on recursive resolvers that support ECS:

A. Encrypt DNS Traffic

DNS over TLS[6] uses Transport Layer Security (TLS) for encrypting DNS traffic. DNSCurve[7] and DNSCrypt[8] use public key cryptography to provide authentication and encryption between caches and servers. Confidential DNS[9] use "ENCRYPT" record to get the public key of the DNS server, and then set the session key to encrypt DNS data.

DNS over TLS, DNSCurve, DNSCrypt and Confidential DNS can improve the privacy on the resolution path, but don't have more influence on the name servers.

B. Reduce Information Leakage to DNS Server

RFC7706[10] describes a method called "Root loopback" for recursive to create an up-to-date root zone server on loopback, hide recursive queries from root. RFC7816[11] describes a technique called "QNAME minimisation", where the resolver no longer sends the full original QNAME to the upstream name server.

Root loopback and QNAME minimisation can hide domain query information from root and TLD, but they are not designed for client IP privacy.

III. EDNS ISP LOCATION (EIL) EXTENSION

The EDNS ISP Location (EIL) extension proposed in this paper is similar to ECS. But EIL includes the geolocation information of client IP in DNS packets, not client subnet address.

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

Like ECS, the authoritative server could provide a better answer by using precise geolocation of client's IP in EIL.

A. Structure

EIL is structured as follows:

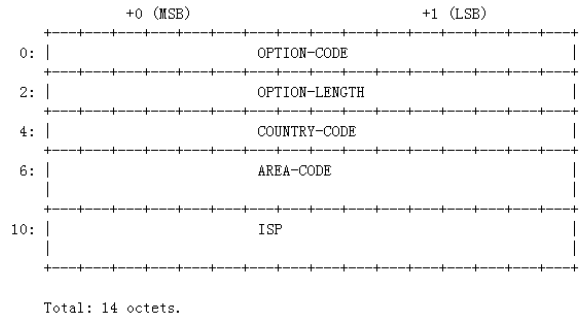

Figure 3. EIL Structure.

1) **OPTION-CODE**, 2 octets, defined in RFC6891[12]. EDNS option code should be assigned by the expert review process as defined by the DNSEXT working group and the IESG.
2) **OPTION-LENGTH**, 2 octets, uppercase, defined in RFC6891, contains the length of the payload (everything after OPTION-LENGTH) in octets.
3) **COUNTRY-CODE**, 2 octets, uppercase, defined in ISO3166[13], indicates the country information of the client's IP. For example, China's COUNTRY-CODE is CN.
4) **AREA-CODE**, 4 octets, uppercase, indicates the area information of the client's IP, using area code of the phone number. AREA-CODE can be found in [14]. For example, Fuzhou is the capital of Fujian Province in China, we can use Fuzhou's area code 0591 to represent the whole Fujian Province.
5) **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-CODE. As Table 1 shows, the shortcut name of China Telecommunications Corporation is TEL.

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 0x20 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*

Discuss AREA-CODE Definition

The AREA-CODE is still part of future research. Take client IP 61.154.123.91 for example.

As Figure 4 shows, Maxmind GeoIP Database[15] gives the <COUNTRY, AREA, ISP> information:

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

ISO 3166-2[13] defines the country subdivisions code. We can map Location "China-Fujian" to ISO 3166-CODE "CN-35". Therefore, if we directly use ISO 3166-2's country subdivision code for EIL, the area precision is on the province level.

If we use area code of the phone number for EIL, the area precision can be raised to the city level and support numeric string fuzzy matching on the 4 octets digital area code. For example, Quanzhou City's area code is 0595, fuzzy matching Fuzhou's 0591.

GeoIP2 City Database Demo

IP Addresses

61.154.123.91

Enter up to 25 IP addresses separated by spaces or commas. You can also test you

Submit

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

Figure 4. Maxmind GeoLocation.

B. Deploy

1) P-model: EIL initiated at public recursive resolver

When a public recursive resolver receives a DNS query from local forwarding resolver, it can guess geolocation of client's IP and generate the EIL OPT data, then send EIL query to the authoritative server.

As Figure 5 shows, this will move the "guess geolocation of client's IP" work from authoritative server to public recursive resolver, lighten the burden of authoritative server, but increase DDoS risk on public recursive resolver.

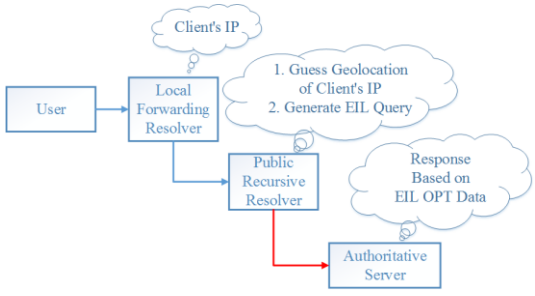

Figure 5. EIL initiated at public recursive resolver.

2) L-model: EIL initiated at local forwarding resolver

Local forwarding resolver is usually on the first-hop router, such as public Wi-Fi hotspot routers and Cisco/Linksys/Netgear/TP-LINK home routers.

As Figure 6 shows, when a local forwarding resolver that implements EIL receives a DNS query from an end user, it surely can know about the geolocation information of client's IP, and generate the EIL OPT data, then send the EIL query to the intermediate recursive resolver. Intermediate recursive resolver sends the EIL query to the authoritative server.

In this condition, both public recursive resolver and authoritative server don't need to "guess geolocation of client's IP", because the local forwarding resolver supplies the geolocation precisely. That is, EIL can reduce dependence on the IP geolocation database quality, which is crucial to DNS response accuracy in ECS.

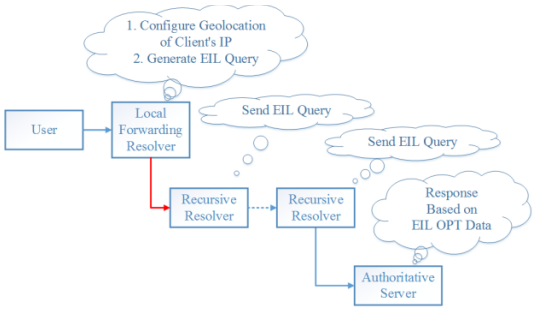

Figure 6. EIL initiated at local forwarding resolver.

3) I-model: EIL initiated at ISP recursive resolver

ISP recursive resolver only serves its customers, each of whom has a static geolocation. As Figure 7 shows, ISP recursive resolver can add EIL transparent to end user, and then authoritative server doesn't need to "guess geolocation of client's IP". EIL will benefit if the authoritative server could not find the approximate geolocation of ISP recursive resolver, which is crucial to DNS response accuracy in ECS.

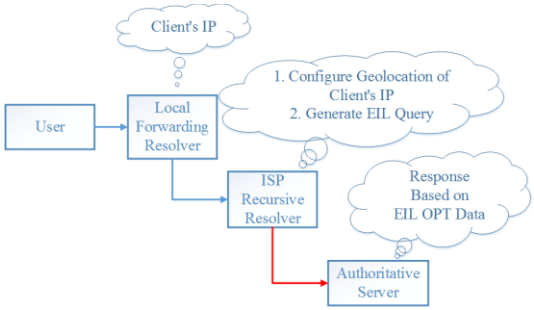

Figure 7. EIL initiated at ISP recursive resolver.

C. Response

Using the geolocation information specified in the EIL of DNS query, the authoritative server can generate a tailored response.

Authoritative servers that not supporting EIL ought to safely ignore it within incoming queries, and response the query as a normal case without EDNS option.

EIL contains a whitelist for COUNTRY-CODE, AREA-CODE and ISP, which can be maintained by the DNSOP working group. Authoritative servers that supporting EIL must only response the EIL queries matched the whitelist. Recursive resolver that supporting EIL must only cache the EIL responses matched the whitelist.

D. Support ECS and EIL at the same time

Name servers can support ECS and EIL at the same time. But ECS and EIL can't be both initiated at the same dns packet. It is better for user privacy if name servers initiate the EIL query prior to the ECS query.

Imagine that authoritative servers support both ECS and EIL. Recursive resolvers can cache both ECS response and EIL response, table 2 shows some choices for recursive resolvers when they receive dns queries.

Table 2. ECS and EIL at Recursive Resolver.

Receive EIL query: *Search in EIL cache.* *If cache is matched, return EIL response.* *Otherwise, send EIL query to authoritative server.* ***Receive ECS query:*** *Search in ECS cache.* *If cache is matched, return ECS response.* *Otherwise, send ECS query to authoritative server.* ***Receive DNS query without EDNS option:*** *Search in ECS cache.* *If cache is matched, return ECS response.* *Otherwise,* *Guess the geolocation information of the client's IP, build EIL option for the query packet.* *Search in EIL cache.* *If cache is matched, return EIL response.* *Otherwise, send EIL query to authoritative server.* ***Receive DNS query with not-ECS/not-EIL option:*** *Search in not-EDNS cache.* *If cache is matched, return response.* *Otherwise, send the DNS query to authoritative server.*
Receive ECS query, improve user privacy: *Guess the geolocation information of the client's IP, build EIL option for the query packet.* *Search in EIL cache.* *If cache is matched, return EIL response RR with origin ECS option.* *Otherwise, send EIL query to authoritative server.*

E. Trade-off of the three deployment models

As described above, EIL has three deployment models: P-model, L-model and I-model. Table 3 shows a trade-off of the three deployment models.

P-model is the most recommended and close to ECS.

I-model is second recommended, it can reduce the "guess geolocation of client's IP" cost of authoritative servers.

L-model requires firmware upgrade EIL support on the first-hop router, it can reduce the "guess geolocation of client's IP" cost of name servers in the long term.

Table 3. Trade-off of the three deployment models.

Model	*Cost*	*Geolocation quality*	*Deployment Effort*
P-model	*High*	*Depend on GeoIP database*	*Low*
L-model	*Low*	*High*	*High*
I-model	*Low*	*High*	*Middle*

IV. PRIVACY AND SECURITY CONCERNS

A. User Privacy

The biggest privacy concern on ECS is that client subnet information is personally identifiable. The more domains publish their zones on a third-party authoritative server, the more end user privacy information can be gathered by the authoritative server according to the ECS queries.

EIL is aimed to preserve the goodness on end-user experience optimization of ECS, and adjust the sensitive client subnet information to aerial view geolocation information for user privacy protection. Besides, as Figure 6 shows, users can hide themselves from all intermediate recursive resolvers and authoritative servers except the next-hop recursive resolver.

B. Target Censorship

DNS traffic is plain text by default. It is easily to be blocked or poisoned by internet target censorship. To bypass the censorship, it is better to encrypt the dns traffic or use some proxy tunnel.

EIL's geolocation information covers bigger area than ECS's client subnet information. Therefore, compared to ECS in plain text condition, EIL is weaker at blocking record attack, but stronger at targeted DNS poisoning attack.

C. Public Recursive Cache Size

Like ECS, cache size will raise if a public recursive resolver supports EIL. The cache size of ECS grows up with the number of client subnets. The cache size of EIL is related to the row count in the <COUNTRY-CODE, AREA-CODE, ISP> geolocation whitelist. Therefore, under IPv6 environment, the cache size of EIL will be smaller than ECS.

Table 4. Geolocation In China.

Geolocation Type	*Configuration Number*
Area + ISP	*(23+11)*5 = 170*
Area + NULL ISP	*(23+11)*1 = 34*
NULL AREA + ISP	*1*5 = 5*
NULL AREA + NULL ISP	*1*1 = 1*
Total	*170+34+5+1=210*

Let's take the example of China. There are 23 provinces and 11 special zones in China. As Table 1

shows, TEL, UNI, MOB, TIE and EDU are the top 5 ISP in China. As Table 4 shows, consider the null value of AREA-CODE and ISP, there will be 210 configurations on the authoritative server to match the geolocation of China. This is the maximum cache size of EIL on public recursive resolver for China.

D. DDoS Attack

Name servers optional only implement EIL query when the query is from a TCP connection to defense spoofed IP addresses.

Like pseudo-random sub-domain attack, name servers 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.

Table 5 shows the pseudocode to deal with error EIL query in Perl style. Authoritative server should refuse all error EIL query for security (strict_eil). Recursive resolver could choose to make a better EIL query instead of refusing if it thinks itself can afford (guess_eil).

Table 5. Deal With Error EIL Query.

```
our %EIL_W;  # whitelist

sub strict_eil {
    my ($eil) = @_;
    my ($c, $a, $i) =
     @{$eil}{'country_code','area_code','isp'};

    return unless(exists $EIL_W{$c});
    return unless(exists $EIL_W{$c}{area}{$a});
   return unless(exists $EIL_W{$c}{isp}{$i});

    return $eil;
}


sub guess_eil {
     my ($eil) = @_;

    my ($c, $a, $i) =
        @{$eil}{'country_code','area_code','isp'};

    $eil->{'country_code'}='' unless(exists $EIL_W{$c});

   #area_code will be null if country_code is null
   $eil->{'area_code'}=''
       unless(exists $EIL_W{$c}{area}{$a});

   $eil->{'isp'}='' unless(exists $EIL_W{$c}{isp}{$i});

    return $eil;
}
```

V. CONCLUSION

It is unrealistic to deny the internet content delivery acceleration brought by ECS because of the privacy concerns. The goal of EIL is to preserve the end-user experience optimization and make some privacy improvement on ECS. Note that name servers should only enable EIL where it is expected to benefit the end users, such as dealing with some latency-sensitive CDN domain queries in a complex network environment. We believe that EIL can provide user privacy controls on public recursive resolvers and authoritative servers.

Our demo codes of EIL can be found in Github: https://github.com/abbypan/dns_test_eil. Our future work is to do more experiments in China network environment. We wish to apply the EIL to real DNS traffic in the future and bring it to the IETF.

REFERENCES

[1] Sajal. "Which CDNs support edns-client-subnet?" <http://www.cdnplanet.com/blog/which-cdns-support-edns-client-subnet/>, 2012.

[2] Contavalli, C., W. van der Gaast, and D. Lawrence. W. Kumari," Client Subnet in DNS Queries. RFC 7871, DOI 10.17487/RFC7871, May 2016,< http://www. rfc-editor. org/info/rfc7871>.

[3] Kintis, Panagiotis, et al. "Understanding the Privacy Implications of ECS." Detection of Intrusions and Malware, and Vulnerability Assessment. Springer International Publishing, 2016. 343-353.

[4] Bortzmeyer, Stephane. DNS privacy considerations. No. RFC 7626. 2015.

[5] Grothoff, Christian, and Matthias Wachs Monika Ermert Jacob Appelbaum. "NSA's MORECOWBELL: Knell for DNS."

[6] Hu, Z., et al. Specification for DNS over Transport Layer Security (TLS). No. RFC 7858. 2016.

[7] Dempsky, M. "Dnscurve: Link-level security for the domain name system." Work in Progress, draft-dempsky-dnscurve-01 (2010).

[8] https://dnscrypt.org/

[9] Wijngaards, W., and G. Wiley. "Confidential DNS." IETF Draft.(https://tools.ietf.org/html/draft-wijngaards-dnsop-confidentialdns-03) (2015).

[10] Kumari, W., and P. Hoffman. Decreasing Access Time to Root Servers by Running One on Loopback. No. RFC 7706. 2015.

[11] Bortzmeyer, S. "RFC 7816–DNS query name minimisation to improve privacy." DNS Query Name Minimisation to Improve Privacy (2016).

[12] Damas, Joao, Michael Graff, and Paul Vixie. Extension mechanisms for DNS (EDNS (0)). No. RFC 6891. 2013.

[13] ISO 3166, "Country Codes", <http://www.iso.org/iso/country_codes>.

[14] Fusion Labs, "Area Codes", <http://www.area-codes-db.com/>.

[15] Maxmind, "GeoIP2 City Database", <https://www.maxmind.com/en/geoip-demo>.