Deep Dive into NXDOMAIN Data in China

Jinghua Bai, Zaifeng Zhang

DNS-OARC 42 February 2024

Contents

Introduction

QAX operates a biggest public DNS resolver in China:

- Trillions of queries per day
- Cover 34 provinces
- Largest public PassiveDNS system

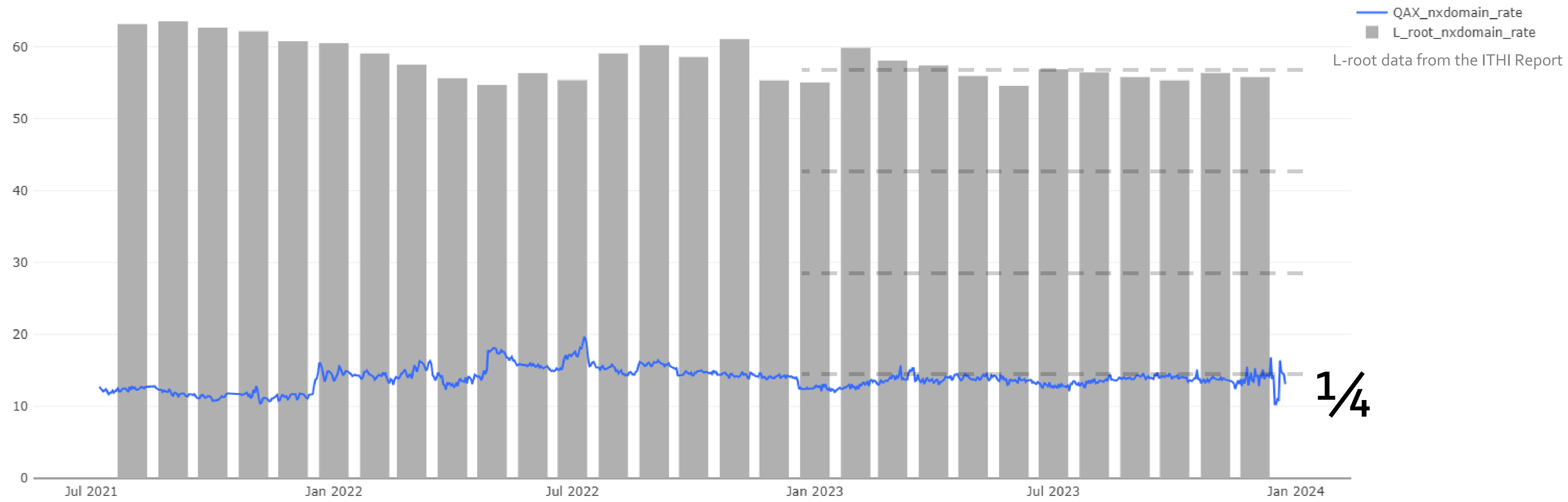
We have built a system to monitor NXDOMAIN data from various perspectives:

- NXDOMAIN proportion
- TLD distribution
- Domain patterns analysis
- Client regional analysis

I am going to show the results and some findings.

CONTENTS

NXDOMAIN response rate

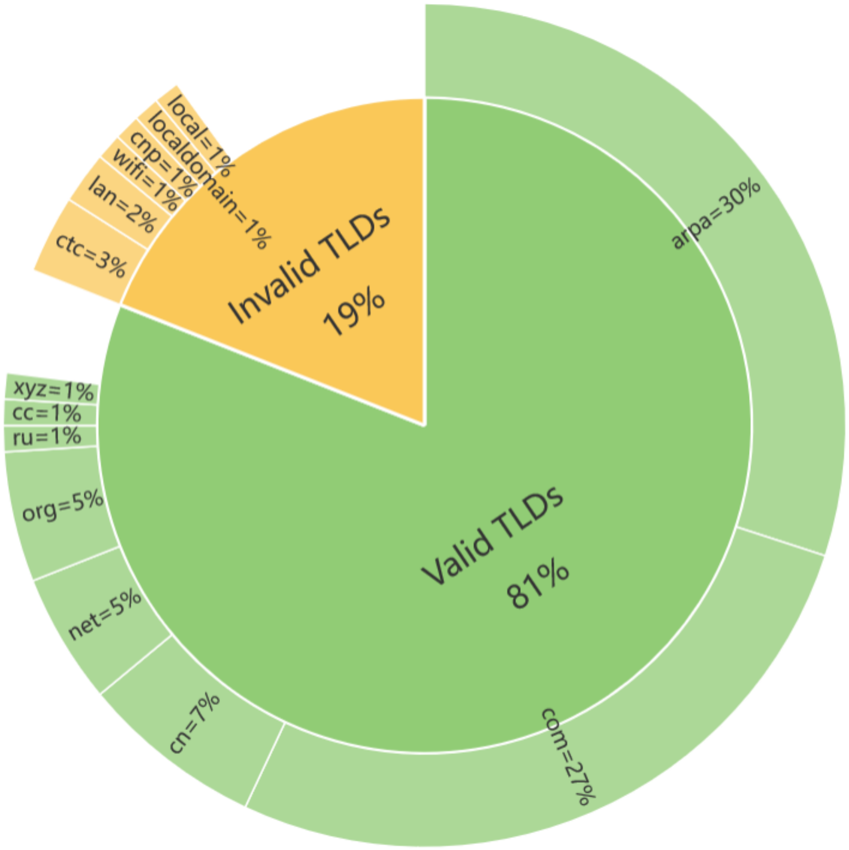


The NXDOMAIN response rate of the QAX recursive resolver is a quarter of that of the L-root server.

- Recursive resolvers face the users directly and often encounter repeated queries.
- Root servers mainly handle top-level domains (TLDs) and are more likely to encounter non-existent domain names.

Valid TLDs (ICANN) vs. Invalid TLDs (non-ICANN)

NXDOMAIN in recursive resolvers indicates that the domain name does not exist, not that the TLD does not exist.



81% Valid TLDs

- .arpa: over half of the queries are PTR queries for private IP addresses
- .cn: China's ccTLD

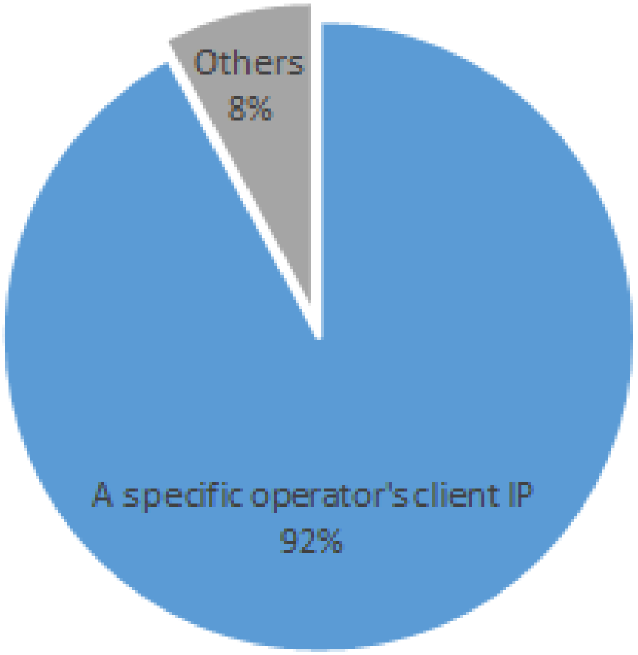
19% Invalid TLDs

- .ctc: a special suffix used by a Chinese telecom operator

Invalid TLDs (non-ICANN)

RANK	NON-ICANN TLD	% OF NON-ICANN	QUANTITY SCALE/DAILY	FQDN QUANTITY SCALE/DAILY	RANK IN L-ROOT SERVER
1	ctc	16.73%	1B+	10M+	7
2	lan	8.07%	1B+	10M+	5
3	wifi	4.50%	100M+	1M+	12
4	cnp	4.23%	100M+	<1k	20+
5	localdomain	3.55%	100M+	100k+	11
6	local	3.11%	100M+	10M+	1
7	dhcp	2.13%	100M+	1M+	4
8	3132372e302e302e31	2.08%	100M+	1	20+
9	novalocal	2.07%	100M+	1M+	20+
10	rl=http	1.97%	100M+	1	20+
11	comp	1.50%	100M+	<1k	20+
12	openstacklocal	1.46%	100M+	1M+	20
13	0	1.38%	100M+	10k+	20+
14	localhost	1.33%	100M+	100k+	10
15	home	1.13%	100M+	1M+	3
16	***-wlan-controller	1.00%	100M+	1	20+
17	url	0.74%	100M+	<1k	20+
18	br-lan	0.62%	100M+	1	20+
19	bbrouter	0.55%	100M+	100k+	8
20	null	0.49%	10M+	10k+	20+

Case1: ctc is a special suffix used by a Chinese telecom operator



1. Most client IP addresses querying '.ctc' domains (92%) are from a specific Chinese telecom operator.
2. The standard login address for the customized routers provided by this operator is 'router.ctc'.
3. The acronym for this telecom operator is 'CTC'.

Case2: cnp

CNP issue: Incorrect 'p' added to a cloud service provider's subdomains, causing numerous NXDOMAIN errors.

.CNP FQDN	% OF .CNP QUERIES	DAILY QUERIES QUANTITY
n-relay-ipc-txc-nj-00.***cloud.com.cnp	62.58%	100M+
n-txc-relay-ipc-nj-01.***cloud.com.cnp	24.13%	100M+
n-txc-relay-ipc-nj-00.***cloud.com.cnp	12.23%	10M+

98%

OpenNIC, Tor, and Namecoin in Public DNS.

TAG	TLD	% OF NON-ICANN NXDOMAIN QUERIES	DAILY QUERIES QUANTITY SCALE	DAILY FQDN QUANTITY SCALE
OpenNIC	null	0.4914%	10M+	10k+
Tor	onion	0.0054%	1M+	1k+
OpenNIC	o	0.0014%	100k+	1k+
Namecoin	bit	0.0002%	10k+	10+
OpenNIC	oss	0.0002%	10k+	10+
OpenNIC	pirate	0.0001%	10k+	<10
OpenNIC	geek	0.0001%	10k+	10+
OpenNIC	libre	0.0001%	10k+	10+
OpenNIC	gopher	0.0001%	10k+	<10
OpenNIC	bbs	0.0000%	1k+	10+
OpenNIC	parody	0.0000%	1k+	<10
OpenNIC	dyn	0.0000%	1k+	10+
OpenNIC	indy	0.0000%	1k+	<10
OpenNIC	chan	0.0000%	<1k	100+
OpenNIC	oz	0.0000%	<1k	100+
OpenNIC	cyb	0.0000%	<1k	10+
OpenNIC	neo	0.0000%	<1k	10+

Reasons for leaking to public DNS:

- Erroneous configurations (such as some test sandboxes).
- Incorrect usage (like non-Tor browsers attempting to connect to Tor domain names).

Contents

Purpose of analyzing domain name patterns

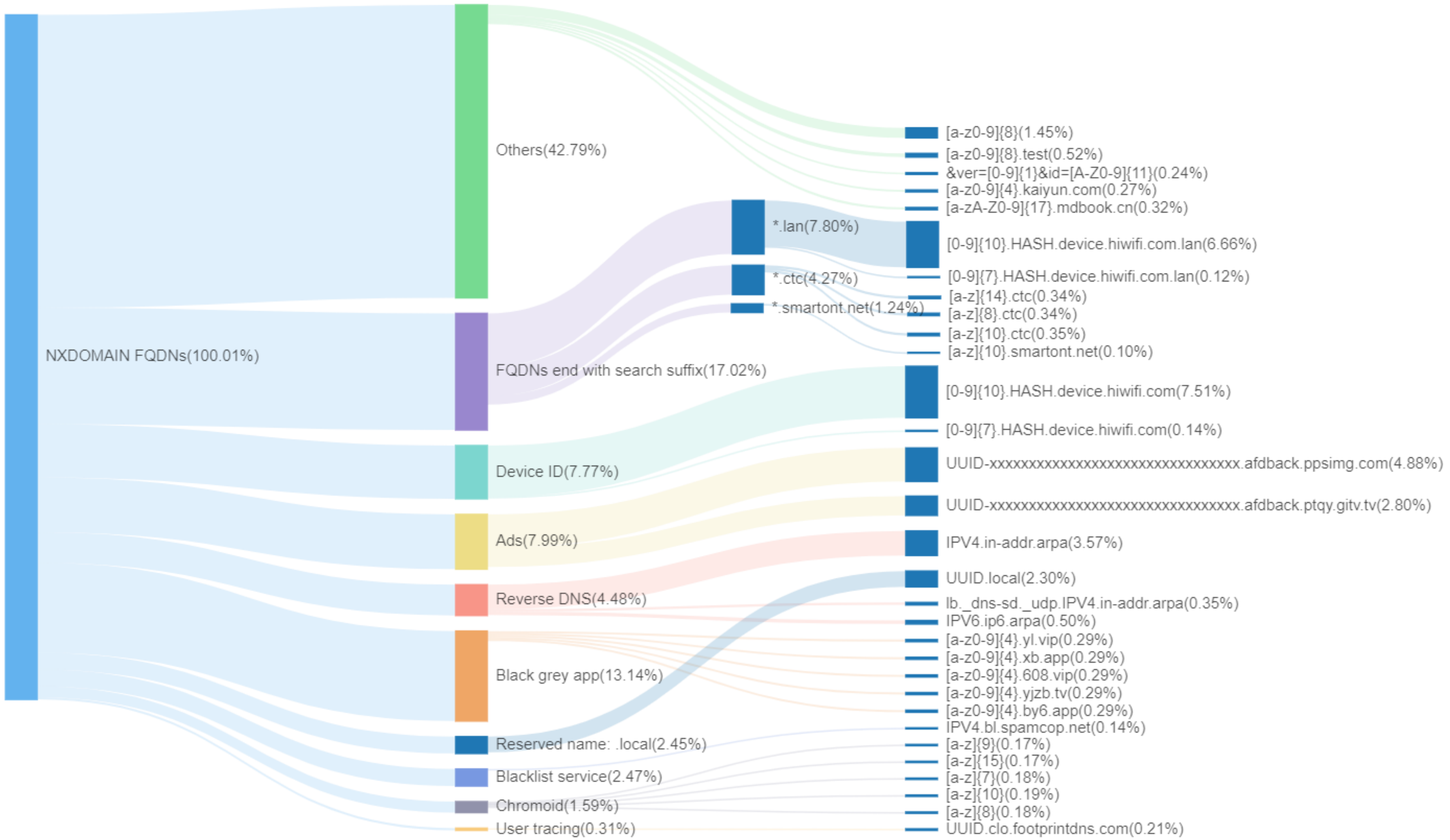
- Understanding Domain Characteristics: Conduct in-depth analysis of domain structures to comprehensively understand the features and underlying reasons of domain names in the network.
- Identifying Anomalous Network Behavior: Analyze NXDOMAIN response patterns to reveal abnormal activities triggered by specific domain names.
- Optimizing DNS Server Performance: By identifying abnormal behaviors through domain pattern analysis, DNS server load can be reduced through targeted optimizations.

Prominent domain patterns in NXDOMAIN

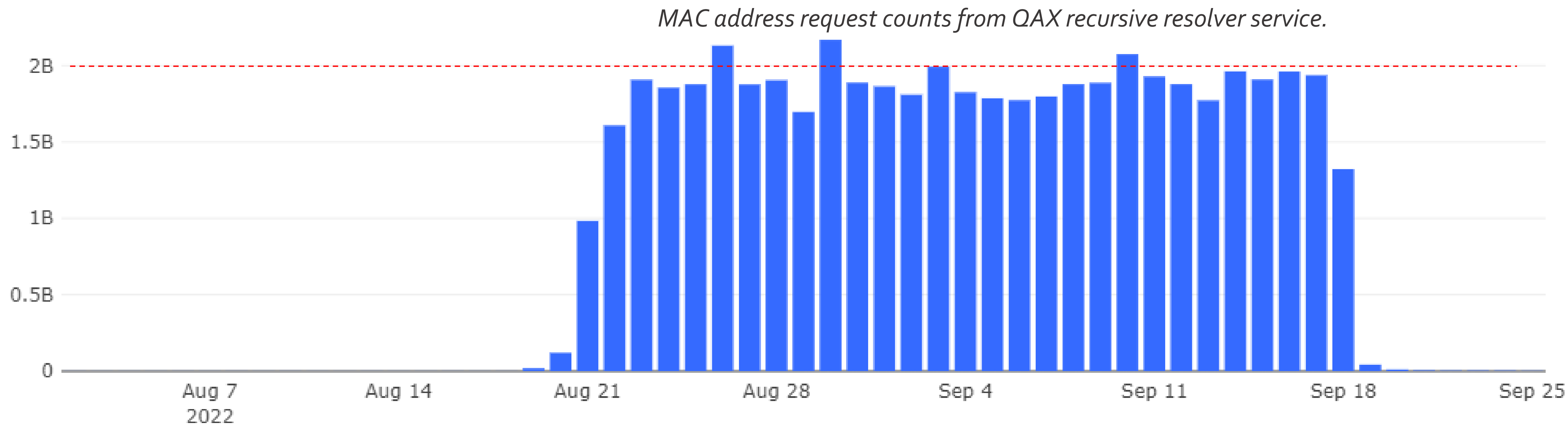
We monitored NXDOMAIN domain names for nearly 5 months. Hundreds of millions of unique domain names are monitored everyday.

Classification Results:

1. software: Black grey app, Ads(advertisements), Blacklist service, Chromoid, User tracing
2. system configuration: FQDNs end with search suffix, Device ID, Reverse DNS



Case1： High volume of MAC address DNS queries from a top-level APP



1. In September 2022, we received a lead from a root server operation team, who discovered a large number of DNS queries in the form of MAC addresses coming from China on their root servers.
2. Further analysis confirmed that this phenomenon was related to the Top-level app in our country.
3. We subsequently reported this bug to the relevant departments, and after the bug was fixed on September 18th, such requests no longer appeared.

Case1： High volume of MAC address DNS queries from a top-level APP (Continue)

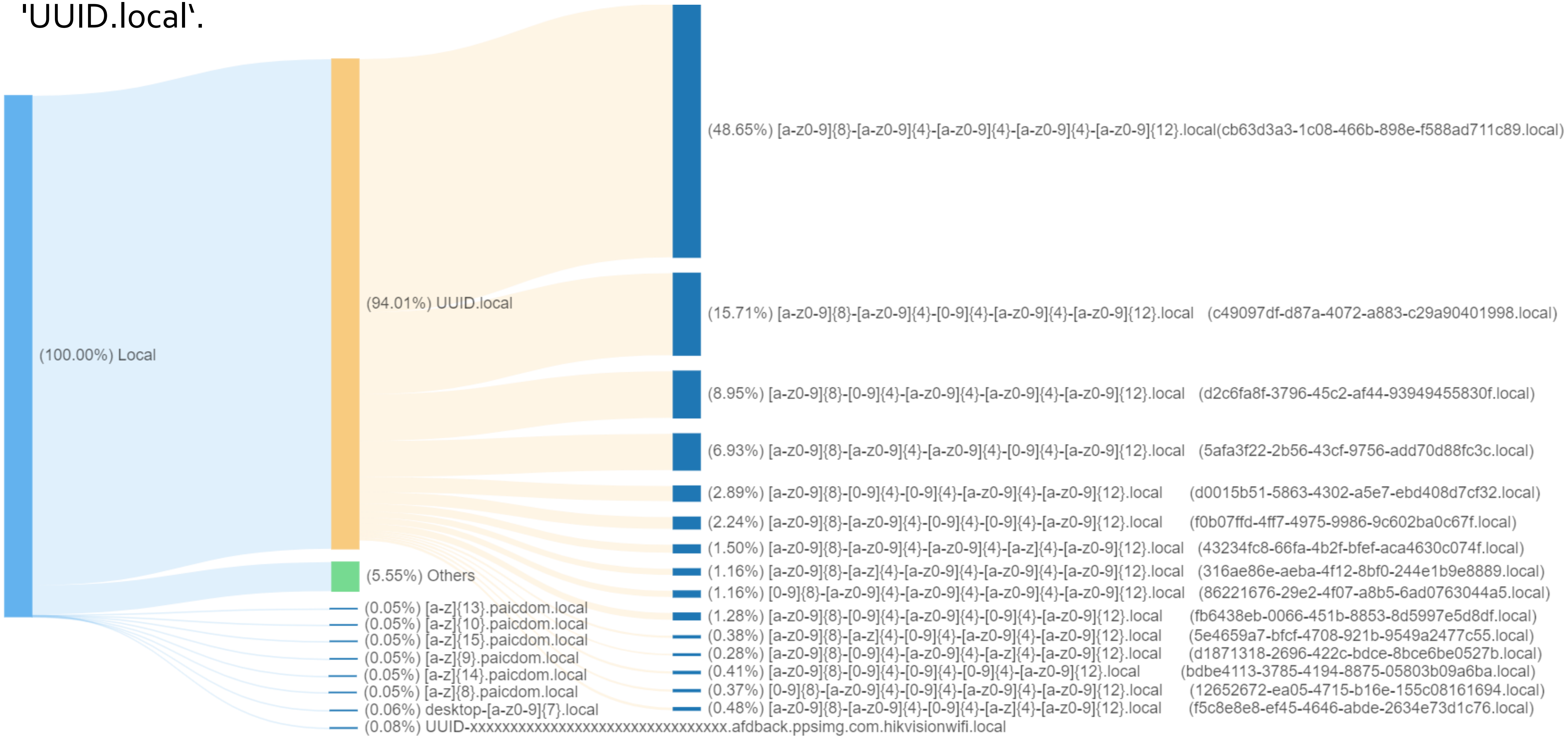
```
Standard query 0x9617 A dev
Standard query 0x9617 A dev
Standard query response 0x9617 Server failure A dev
Standard query response 0x9617 Server failure A dev
Standard query 0x9089 A wlan0
Standard query 0x9089 A wlan0
Standard query response 0x9089 Server failure A wlan0
Standard query response 0x9089 Server failure A wlan0
Standard query 0xb2bd A lladdr
Standard query 0xb2bd A lladdr
Standard query response 0xb2bd Server failure A lladdr
Standard query response 0xb2bd Server failure A lladdr
Standard query 0xda03 A 64:2f:c7:a1:fc:01
Standard query 0xda03 A 64:2f:c7:a1:fc:01
Standard query response 0xda03 Server failure A 64:2f:c7:a1:fc:01
Standard query response 0xda03 Server failure A 64:2f:c7:a1:fc:01
Standard query 0xd20b A REACHABLE
Standard query 0xd20b A REACHABLE
Standard query response 0xd20b Server failure A REACHABLE
Standard query response 0xd20b Server failure A REACHABLE
```

```
C:\Windows\System32>adb shell ip neigh show
192.168.43.181 dev wlan0 lladdr 20:f4:78:09:48:b3 DELAY
fe80::22f4:78ff:fe09:48b3 dev wlan0 lladdr 20:f4:78:09:48:b3 DELAY
2409:894c:130:10bc:d:1b8e:e8ea:3b2c dev wlan0 lladdr 20:f4:78:09:48:b3 REACHABLE
```

Bug detail: A function of the application needs to process the return result when checking the Wi-Fi hotspot link. The result program has a bug that causes the DNS query of the return result, which contains the MAC address used.

Case2: UUID.local

We have discovered that domain names ending with 'local' almost exclusively appear in the form of 'UUID.local‵.



Case2: How is UUID.local generated?

WebRTC is a free, open-source project. It lets web browsers and mobile applications add capabilities for real-time audio and video directly between users. However, WebRTC has the risk of leaking the user's real IP address, both public and private.

IETF draft proposes using dynamic mDNS names (UUID.local) to conceal private IP addresses, such as the 'Anonymize local IPs exposed by WebRTC' option in Chrome settings.

Expires: 8 June 2022 Q. Wang
Google

Using Multicast DNS to protect privacy when exposing ICE candidates

Abstract

WebRTC applications collect ICE candidates as part of the process of creating peer-to-peer connections. To maximize the probability of a direct peer-to-peer connection, the endpoint's local IP addresses are included in this candidate collection. However, these addresses are typically private and, as such, their disclosure has privacy implications. This document describes a privacy-preserving way to share local IP addresses with other endpoints by concealing the addresses with dynamically generated Multicast DNS (mDNS) names.

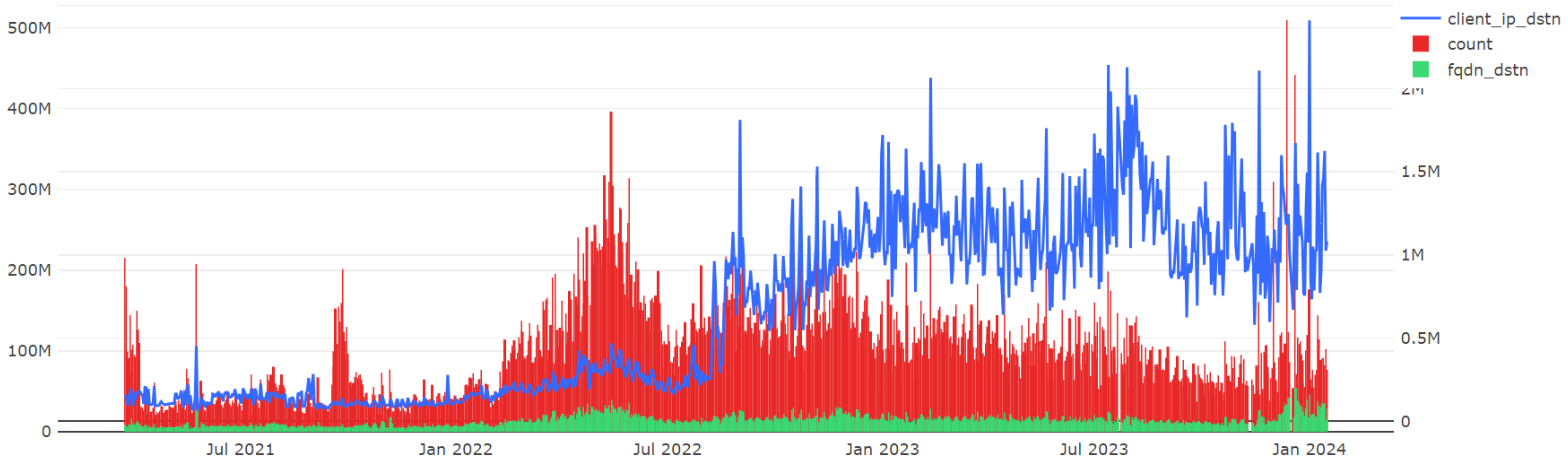
Source: https://datatracker.ietf.org/doc/html/draft-ietf-mmusic-mdns-ice-candidates

Case2: How is UUID.local generated? (Continue)

- mDNS itself is harmless and does not leak into the public DNS environment, but such a large number of domain name queries for UUID.local may be related to configuration or application bugs.
- When using a specific video application on an Android phone and clicking on the "Live" interface, it triggers domain name queries in the form of UUID.local.
- The L root server had the highest number of queries for .local, which may indicate widespread global use of WebRTC and similar bugs. (Source: https://ithi.research.icann.org/graph-m3.html)

5518	62.433087	192.168.2.100	192.168.1.1	DNS	105	Standard query 0x7992 A cn-sdqd-ccc-live-tracker-04.chat.[redacted]
5519	62.447098	192.168.1.1	192.168.2.100	DNS	121	Standard query response 0x7992 A cn-sdqd-ccc-live-tracker-04.chat.[redacted] A 120.2[redacted]
5614	64.423628	192.168.2.100	192.168.1.1	DNS	84	Standard query 0x22da A stun-1.chat.[redacted]m
5616	64.437280	192.168.1.1	192.168.2.100	DNS	100	Standard query response 0x22da A stun-1.chat.[redacted]m A 120.[redacted]
5617	64.445693	192.168.2.100	192.168.1.1	DNS	84	Standard query 0x6691 A stun-2.chat.[redacted]
5618	64.456442	192.168.1.1	192.168.2.100	DNS	100	Standard query response 0x6691 A stun-2.chat.[redacted]m A 120.[redacted]
5619	64.467066	192.168.2.100	192.168.1.1	DNS	84	Standard query 0x4299 A stun-3.chat.[redacted]
5620	64.479278	192.168.1.1	192.168.2.100	DNS	100	Standard query response 0x4299 A stun-3.chat.[redacted] A 120.[redacted]
5622	64.482726	192.168.2.100	192.168.1.1	DNS	84	Standard query 0x21d3 A stun-4.chat.[redacted]
5623	64.494033	192.168.1.1	192.168.2.100	DNS	100	Standard query response 0x21d3 A stun-4.chat.[redacted] A 120.[redacted]
5643	64.558067	192.168.2.100	192.168.1.1	DNS	102	Standard query 0x4c37 A 1390c717-874b-41ec-bf84-4c52f09051cd.local
5696	64.569692	192.168.1.1	192.168.2.100	DNS	102	Standard query response 0x4c37 No such name A 1390c717-874b-41ec-bf84-4c52f09051cd.local
5713	64.574287	192.168.2.100	192.168.1.1	DNS	102	Standard query 0x7055 A 1390c717-874b-41ec-bf84-4c52f09051cd.local
5732	64.600087	192.168.1.1	192.168.2.100	DNS	102	Standard query response 0x7055 No such name A 1390c717-874b-41ec-bf84-4c52f09051cd.local
5733	64.600948	192.168.2.100	192.168.1.1	DNS	102	Standard query 0x7119 A 1390c717-874b-41ec-bf84-4c52f09051cd.local

Case2: UUID.local trend in QAX recursive resolver



The chart above shows the trend for UUID.local. Although we are unable to capture its initial point, we can observe the scale of its changes through the QAX recursive resolver.

(We note that the largest NXDOMAIN query in the L-root server is for "local")

We will continue to monitor and welcome any questions you may have.

E-mail: baijinghua@qianxin.com
Blog Post: https://blog.xlab.qianxin.com/deep-dive-into-nxdomain-data-in-china/
Twitter: @Xlab_qax